

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:36:23 ; Search time 38.0769 Seconds
(without alignments)
52.493 Million cell updates/sec

Title: US-10-006-593-2

Perfect score: 80

Sequence: 1 IEGFTLRQWLARAP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	15	ABP51670	Thrombopoietin (TP
2	80	100.0	18	ABP51675	TPO mimetic antio
3	80	100.0	18	ABP51684	TPO mimetic antio
4	80	100.0	18	ABP51686	TPO mimetic antio
5	80	100.0	18	ABP51687	TPO mimetic antio
6	80	100.0	18	ABP51688	TPO mimetic antio
7	80	100.0	18	ABP51689	TPO mimetic antio
8	80	100.0	18	ABP51690	TPO mimetic antio
9	80	100.0	18	ABP51691	TPO mimetic antio
10	80	100.0	18	ABP51693	TPO mimetic antio

11	80	100.0	472	23	ABP51695	SG1.1-TPO heavy ch
12	73	91.2	14	18	AAW36774	Thrombopoietin rec
13	73	91.2	14	18	AAW33030	Thrombopoietin rec
14	73	91.2	14	18	AAW33034	Thrombopoietin rec
15	73	91.2	14	18	AAW09468	Thrombopoietin rec
16	73	91.2	14	18	AAW09463	Thrombopoietin rec
17	73	91.2	14	21	AAE16962	TPO-mimetic peptid
18	73	91.2	14	21	AAE16962	Thrombopoietin mim
19	73	91.2	14	22	AAU25827	Human thrombopoiet
20	73	91.2	14	22	AAU26004	Human thrombopoiet
21	73	91.2	14	23	ABP51669	Thrombopoietin (TP
22	73	91.2	14	23	AAE18011	Human ligand #31 a
23	73	91.2	14	23	ABE72853	TPO mimetic peptid
24	73	91.2	15	18	AAW36776	Thrombopoietin rec
25	73	91.2	15	18	AAW35416	Thrombopoietin rec
26	73	91.2	15	19	AAW66712	Peptide chain of c
27	73	91.2	15	21	AAW20684	Thrombocyte genera
28	73	91.2	15	22	AAU25831	Human thrombopoiet
29	73	91.2	15	22	AAU25996	Human thrombopoiet
30	73	91.2	16	18	AAW36775	Thrombopoietin rec
31	73	91.2	16	18	AAW36771	Thrombopoietin rec
32	73	91.2	16	18	AAW33035	Thrombopoietin rec
33	73	91.2	16	18	AAW19534	Thrombopoietin rec
34	73	91.2	16	19	AAW66713	Peptide chain of c
35	73	91.2	16	19	AAW66716	Peptide chain of c
36	73	91.2	16	19	AAW66709	Peptide chain of c
37	73	91.2	16	19	AAW66733	Peptide chain of c
38	73	91.2	16	22	AAU25832	Human thrombopoiet
39	73	91.2	16	22	AAU26005	Human thrombopoiet
40	73	91.2	16	22	AAU26043	Human thrombopoiet
41	73	91.2	18	21	AAE16956	PEGylated peptide
42	73	91.2	18	21	AAE16957	PEGylated peptide
43	73	91.2	18	23	ABP51673	TPO mimetic compri
44	73	91.2	18	23	ABP51674	TPO mimetic compri
45	73	91.2	18	23	ABP51677	TPO mimetic peptid

ALIGNMENTS

RESULT 1
ABP51670
ID ABP51670 standard; Peptide; 15 AA.
XX
AC
ABP51670;
XX
DT 01-OCT-2002 (first entry)
XX
Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:2.
DE
XX
XX TPO; EPO; Thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianaemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
XX
XX Homo sapiens.
OS
OS Synthetic.
XX
PN WO200246238-A2.
XX
PD 13-JUN-2002.
XX
XX 05-DEC-2001; 2001WO-US47656.
PF
PR 05-DEC-2000; 2000US-251448P.
PR 04-MAY-2001; 2001US-288889P.
PR 29-MAY-2001; 2001US-294068P.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
XX Bowdish KS, Barbas-Frederickson S, Renshaw M;
XX WPI; 2002-566610/60.
XX

PT A novel immunogen molecule comprising a region in which amino acid
PT residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
PT thrombopoietin mimetic -

XX Claim 19; Page 6; 113pp; English.

CC The present invention describes an immunoglobulin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease.
CC disorders or treatments related to the suppression of haematopoiesis.
CC AB073288 to AB073377, and ABP51669 to ABP51696 represent sequences used
CC in the exemplification of the present invention.

XX Sequence 15 AA;

Query Match 100.0%; Score 80; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGETLQWLAAARAP 15
| | | | | | | | | | | | | | | | |
DB 1 IEGETLQWLAAARAP 15

RESULT 2

ABP51675
ID ABP51675 standard; Peptide; 18 AA.

XX ABP51675;

XX 01-OCT-2002 (first entry)

XX TPO mimetic antibody related peptide graft SEQ ID NO:66.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

OS Synthetic.

XX WO200246238-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US47656.

XX 05-DEC-2000; 2000US-251448P.

PR 04-MAY-2001; 2001US-288889P.

PR 29-MAY-2001; 2001US-294068P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX WPI; 2002-566610/60.

XX
PT
PT
PT
XX
XX
PS

Example 4; Page 55; 113pp; English.

CC The present invention describes an immunoglobulin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease.
CC disorders or treatments related to the suppression of haematopoiesis.
CC AB073288 to AB073377, and ABP51669 to ABP51696 represent sequences used
CC in the exemplification of the present invention.

XX Sequence 18 AA;

Query Match 100.0%; Score 80; DB 23; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGETLQWLAAARAP 15
| | | | | | | | | | | | | | | | |
DB 3 IEGETLQWLAAARAP 17

RESULT 3

ABP51684
ID ABP51684 standard; Peptide; 18 AA.

XX ABP51684;

XX 01-OCT-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:31.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

OS Synthetic.

XX WO200246238-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US47656.

XX 05-DEC-2000; 2000US-251448P.

PR 04-MAY-2001; 2001US-288889P.

PR 29-MAY-2001; 2001US-294068P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;

DR WPI; 2002-566610/60.
 DR N-PSDB; ABQ73362.
 XX
 PT A novel immunogen molecule comprising a region in which amino acid
 PT residues corresponding to at least a portion of the complementary
 PT determining region are replaced or fused with an erythropoietin or
 PT thrombopoietin mimetic -
 XX
 PS Claim 20; Fig 5; 113pp; English.
 XX
 CC The present invention describes an immunoglobulin molecule or its fragment
 CC (I) comprising a region where amino acid residues corresponding to at
 CC least a portion of the complementary determining region (CDR) are
 CC replaced or fused with biologically active peptides e.g. a peptide
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
 CC that is flanked with proline at its carboxy terminus. (I) has
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as
 CC a stimulator of proliferation, differentiation and maturation of
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
 CC for stimulating proliferation, differentiation or growth of
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with
 CC production. (I) with a region where amino acid residues corresponding to
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
 CC production of red blood cells, where (I) is contacted with haematopoietic
 CC stem cells or their progenitors. (I) is useful for diagnostics or
 CC therapeutics, in cell isolation strategies, and for treating patients
 CC suffering from deficiency in cell populations caused by disease,
 CC disorders or treatments related to the suppression of haematopoiesis.
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 18 AA;
 Query Match 100.0%; Score 80; DB 23; Length 18;
 Best Local Similarity 100.0%; Pred. No. 9.9e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLRQWLAAARAP 15
 Db 3 IEGPTLRQWLAAARAP 17
 RESULT 4
 ABP51686
 AC ABP51686 standard; Peptide; 18 AA.
 AC ABP51686;
 DT 01-OCT-2002 (first entry)
 DE TPO mimetic peptide SEQ ID NO:35.
 XX
 KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
 KW complementarity determining region; immunoglobulin; antianaemic;
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200246238-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 05-DEC-2001; 2001WO-US47656.
 XX
 PR 05-DEC-2000; 2000US-251448P.
 PR 04-MAY-2001; 2001US-288889P.
 PR 29-MAY-2001; 2001US-294068P.
 XX
 XX (ALEX-) ALEXION PHARM INC.
 PA
 XX

PI Bowdish KS, Barbas-Frederickson S, Renshaw M;
 XX WPI; 2002-566610/60.
 XX N-PSDB; ABQ73364.
 XX
 PT A novel immunogen molecule comprising a region in which amino acid
 PT residues corresponding to at least a portion of the complementary
 PT determining region are replaced or fused with an erythropoietin or
 PT thrombopoietin mimetic -
 XX
 PS Claim 20; Fig 5; 113pp; English.
 XX
 CC The present invention describes an immunoglobulin molecule or its fragment
 CC (I) comprising a region where amino acid residues corresponding to at
 CC least a portion of the complementary determining region (CDR) are
 CC replaced or fused with biologically active peptides e.g. a peptide
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
 CC that is flanked with proline at its carboxy terminus. (I) has
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as
 CC a stimulator of proliferation, differentiation and maturation of
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
 CC for stimulating proliferation, differentiation or growth of
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with
 CC production. (I) with a region where amino acid residues corresponding to
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
 CC production of red blood cells, where (I) is contacted with haematopoietic
 CC stem cells or their progenitors. (I) is useful for diagnostics or
 CC therapeutics, in cell isolation strategies, and for treating patients
 CC suffering from deficiency in cell populations caused by disease,
 CC disorders or treatments related to the suppression of haematopoiesis.
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 18 AA;
 Query Match 100.0%; Score 80; DB 23; Length 18;
 Best Local Similarity 100.0%; Pred. No. 9.9e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLRQWLAAARAP 15
 Db 3 IEGPTLRQWLAAARAP 17
 RESULT 5
 ABP51687
 ID ABP51687 standard; Peptide; 18 AA.
 AC ABP51687;
 DT 01-OCT-2002 (first entry)
 DE TPO mimetic peptide SEQ ID NO:37.
 XX
 KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
 KW complementarity determining region; immunoglobulin; antianaemic;
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200246238-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 05-DEC-2001; 2001WO-US47656.
 XX
 PR 05-DEC-2000; 2000US-251448P.
 PR 04-MAY-2001; 2001US-288889P.
 PR 29-MAY-2001; 2001US-294068P.
 XX
 XX

PA (ALEX-) ALEXION PHARM INC.
XX Bowdish KS, Barbas-Frederickson S, Renshaw M;
PI WPI; 2002-566610/60.
XX N-PSDB; ABQ73365.
XX A novel immunogen molecule comprising a region in which amino acid
PT residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
XX thrombopoietin mimetic -
PS Claim 20; Fig 5; 113pp; English.
XX The present invention describes an immunoglobulin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianaemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease,
CC disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51699 to ABP51696 represent sequences used
CC in the exemplification of the present invention.
XX SQ Sequence 18 AA;
Query Match 100.0%; Score 80; DB 23; Length 18;
Best Local Similarity 100.0%; Pred. NO. 9.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGETLQWLAAARAP 15
Db 3 IEGETLQWLAAARAP 17
RESULT 6
ABP51688
AC ABP51688 standard; Peptide; 18 AA.
XX AC ABP51688;
XX AC ABP51689;
DT 01-OCT-2002 (first entry)
XX TPO mimetic peptide SEQ ID NO:39.
DE TPO mimetic peptide SEQ ID NO:41.
XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianaemic;
XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
XX Homo sapiens.
OS Synthetic.
XX WO200246238-A2.
XX PN 13-JUN-2002.
XX PD 05-DEC-2001; 2001WO-US47656.
XX PF 05-DEC-2000; 2000US-251448P.
XX PR 04-MAY-2001; 2001US-288889P.
XX PS

PR 29-MAY-2001; 2001US-294068P.
XX (ALEX-) ALEXION PHARM INC.
XX Bowdish KS, Barbas-Frederickson S, Renshaw M;
PI WPI; 2002-566610/60.
XX N-PSDB; ABQ73366.
XX A novel immunogen molecule comprising a region in which amino acid
PT residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
XX thrombopoietin mimetic -
PS Claim 20; Fig 5; 113pp; English.
XX The present invention describes an immunoglobulin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianaemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease,
CC disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51699 to ABP51696 represent sequences used
CC in the exemplification of the present invention.
XX SQ Sequence 18 AA;
Query Match 100.0%; Score 80; DB 23; Length 18;
Best Local Similarity 100.0%; Pred. NO. 9.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGETLQWLAAARAP 15
Db 3 IEGETLQWLAAARAP 17
RESULT 7
ABP51689
ID ABP51689 standard; Peptide; 18 AA.
XX AC ABP51689;
XX AC ABP51689;
DT 01-OCT-2002 (first entry)
XX TPO mimetic peptide SEQ ID NO:41.
DE TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianaemic;
XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
XX Homo sapiens.
OS Synthetic.
XX WO200246238-A2.
XX PN 13-JUN-2002.
XX PD 05-DEC-2001; 2001WO-US47656.
XX PF 05-DEC-2000; 2000US-251448P.
XX PR 04-MAY-2001; 2001US-288889P.

PR 05-DEC-2000; 2000US-251448P.
PR 04-MAY-2001; 2001US-288889P.
XX 29-MAY-2001; 2001US-294068P.
PA (ALEX-) ALEXION PHARM INC.
XX
XX
XX

PI Bowdish KS, Barbas-Frederickson S, Renshaw M;
XX WPI; 2002-566610/60.
DR N-PSDB; ABQ73367.
XX

XX A novel immunogen molecule comprising a region in which amino acid
XX residues corresponding to at least a portion of the complementary
XX determining region are replaced or fused with an erythropoietin or
XX thrombopoietin mimetic -

XX Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment
(I) comprising a region where amino acid residues corresponding to at
least a portion of the complementary determining region (CDR) are
replaced or fused with biologically active peptides e.g. a peptide
mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
that is flanked with proline at its carboxy terminus. (I) has
antianemic, haemostatic and nephrotropic activities, and can be used as
a stimulator of proliferation, differentiation and maturation of
haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
for stimulating proliferation, differentiation or growth of
promegakaryocytes or megakaryocytes, where (I) is contacted with
production. (I) with a region where amino acid residues corresponding to
a portion of CDR is replaced with an EPO mimetic, or which has one or
more of its CDRs fused to an EPO mimetic, is useful for increasing the
production of red blood cells, where (I) is contacted with haematopoietic
stem cells or their progenitors. (I) is useful for diagnostics or
therapeutics, in cell isolation strategies, and for treating patients
suffering from deficiency in cell populations caused by disease,
disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used
in the exemplification of the present invention.

XX Sequence 18 AA;

Query Match 100.0%; Score 80; DB 23; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 IEGETLRQWLARAP 15
| | | | | | | | | | | | | | | | | |
3 IEGETLRQWLARAP 17

RESULT 8
ABP51690
ID ABP51690 standard; Peptide; 18 AA.

AC ABP51690;

DT 01-OCT-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:43.

KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.
OS Synthetic.

XX WO200246238-A2.

XX 13-JUN-2002.

XX

PF 05-DEC-2001; 2001WO-US47656.

XX 05-DEC-2000; 2000US-251448P.

PR 04-MAY-2001; 2001US-288889P.

XX 29-MAY-2001; 2001US-294068P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX WPI; 2002-566610/60.

DR N-PSDB; ABQ73368.

XX A novel immunogen molecule comprising a region in which amino acid
XX residues corresponding to at least a portion of the complementary
XX determining region are replaced or fused with an erythropoietin or
XX thrombopoietin mimetic -

XX Claim 20; Fig 5; 113pp; English.

XX The present invention describes an immunoglobulin molecule or its fragment
(I) comprising a region where amino acid residues corresponding to at
least a portion of the complementary determining region (CDR) are
replaced or fused with biologically active peptides e.g. a peptide
mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
that is flanked with proline at its carboxy terminus. (I) has
antianemic, haemostatic and nephrotropic activities, and can be used as
a stimulator of proliferation, differentiation and maturation of
haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
for stimulating proliferation, differentiation or growth of
promegakaryocytes or megakaryocytes, where (I) is contacted with
production. (I) with a region where amino acid residues corresponding to
a portion of CDR is replaced with an EPO mimetic, or which has one or
more of its CDRs fused to an EPO mimetic, is useful for increasing the
production of red blood cells, where (I) is contacted with haematopoietic
stem cells or their progenitors. (I) is useful for diagnostics or
therapeutics, in cell isolation strategies, and for treating patients
suffering from deficiency in cell populations caused by disease,
disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used
in the exemplification of the present invention.

XX Sequence 18 AA;

Query Match 100.0%; Score 80; DB 23; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGETLRQWLARAP 15
| | | | | | | | | | | | | | | | | |
Db 3 IEGETLRQWLARAP 17

RESULT 9

ABP51691

ID ABP51691 standard; Peptide; 18 AA.

AC ABP51691;

DT 01-OCT-2002 (first entry)

DE TPO mimetic peptide SEQ ID NO:45.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.
OS Synthetic.

XX WO200246238-A2.

XX

13-JUN-2002.

05-DEC-2001; 2001WO-US47656.

05-DEC-2000; 2000US-251448P.

04-MAY-2001; 2001US-288889P.

29-MAY-2001; 2001US-294068P.

(ALEX-) ALEXION PHARM INC.

Bowdish KS, Barbas-Frederickson S, Renshaw M;

WPI; 2002-566610/60.

N-PSDB; ABQ73369.

A novel immunogen molecule comprising a region in which amino acid residues corresponding to at least a portion of the complementary determining region are replaced or fused with an erythropoietin or thrombopoietin mimetic -

Claim 20; Fig 5; 113pp; English.

The present invention describes an immunoglobulin molecule or its fragment (I) comprising a region where amino acid residues corresponding to at least a portion of the complementary determining region (CDR) are replaced or fused with biologically active peptides e.g. a peptide mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic, that is flanked with proline at its carboxy terminus. (I) has an antianaemic, haemostatic and nephrotropic activities, and can be used as a stimulator of proliferation, differentiation and maturation of haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful for stimulating proliferation, differentiation or growth of promegakaryocytes or megakaryocytes, where (I) is contacted with promegakaryocytes or megakaryocytes, which results in increased platelet production. (I) with a region where amino acid residues corresponding to a portion of CDR is replaced with an EPO mimetic, or which has one or more of its CDRs fused to an EPO mimetic, is useful for increasing the production of red blood cells, where (I) is contacted with haematopoietic stem cells or their progenitors. (I) is useful for diagnostics or therapeutics, in cell isolation strategies, and for treating patients suffering from deficiency in cell populations caused by disease, disorders or treatments related to the suppression of haematopoiesis. CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in the exemplification of the present invention.

Sequence 18 AA;

Query Match 100.0%; Score 80; DB 23; Length 18;

Best Local Similarity 100.0%; Pred. No. 9.9e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IEQPTLRQWLARAP 15

3 IEQPTLRQWLARAP 17

RESULT 10

ABP51693

ID ABP51693 standard; Peptide; 18 AA.

XX ABP51693;

XX 01-OCT-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:49.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
 XX complementarity determining region; immunoglobulin; antianaemic;
 XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
 OS Homo sapiens.
 XX Synthetic.

WO200246238-A2.

13-JUN-2002.

05-DEC-2001; 2001WO-US47656.

05-DEC-2000; 2000US-251448P.

04-MAY-2001; 2001US-288889P.

29-MAY-2001; 2001US-294068P.

(ALEX-) ALEXION PHARM INC.

Bowdish KS, Barbas-Frederickson S, Renshaw M;

WPI; 2002-566610/60.

N-PSDB; ABQ73371.

A novel immunogen molecule comprising a region in which amino acid residues corresponding to at least a portion of the complementary determining region are replaced or fused with an erythropoietin or thrombopoietin mimetic -

Claim 20; Fig 5; 113pp; English.

The present invention describes an immunoglobulin molecule or its fragment (I) comprising a region where amino acid residues corresponding to at least a portion of the complementary determining region (CDR) are replaced or fused with biologically active peptides e.g. a peptide mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic, that is flanked with proline at its carboxy terminus. (I) has an antianaemic, haemostatic and nephrotropic activities, and can be used as a stimulator of proliferation, differentiation and maturation of haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful for stimulating proliferation, differentiation or growth of promegakaryocytes or megakaryocytes, where (I) is contacted with promegakaryocytes or megakaryocytes, which results in increased platelet production. (I) with a region where amino acid residues corresponding to a portion of CDR is replaced with an EPO mimetic, or which has one or more of its CDRs fused to an EPO mimetic, is useful for increasing the production of red blood cells, where (I) is contacted with haematopoietic stem cells or their progenitors. (I) is useful for diagnostics or therapeutics, in cell isolation strategies, and for treating patients suffering from deficiency in cell populations caused by disease, disorders or treatments related to the suppression of haematopoiesis. CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in the exemplification of the present invention.

Sequence 18 AA;

Query Match 100.0%; Score 80; DB 23; Length 18;

Best Local Similarity 100.0%; Pred. No. 9.9e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IEQPTLRQWLARAP 15

3 IEQPTLRQWLARAP 17

RESULT 11

ABP51695

ID ABP51695 standard; Protein; 472 AA.

XX ABP51695;

XX 01-OCT-2002 (first entry)

XX 5G1.1-TPO heavy chain amino acid sequence SEQ ID NO:67.

XX TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
 XX complementarity determining region; immunoglobulin; antianaemic;
 XX haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
 OS Homo sapiens.

OS Synthetic.
 PN WO200246238-A2.
 PD XX
 PE 13-JUN-2002.
 PF XX
 PG 05-DEC-2001; 2001WO-US47656.
 PH XX
 PI 05-DEC-2000; 2000US-251448P.
 PJ 04-MAY-2001; 2001US-288889P.
 PK 29-MAY-2001; 2001US-294068P.
 PL XX
 PM (ALEX-) ALEXION PHARM INC.
 PN XX
 PO Bowdish KS, Barbas-Frederickson S, Renshaw M;
 PP WPI; 2002-566610/60.
 PQ N-PSDB; ABQ73374.
 PR A novel immunogen molecule comprising a region in which amino acid
 residues corresponding to at least a portion of the complementary
 determining region are replaced or fused with an erythropoietin or
 thrombopoietin mimetic -
 PT
 PU
 PV
 PW
 PX Example 4; Fig 13A; 113pp; English.
 PY The present invention describes an immunoglobulin molecule or its fragment
 (I) comprising a region where amino acid residues corresponding to at
 least a portion of the complementary determining region (CDR) are
 replaced or fused with biologically active peptides e.g. a peptide
 mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
 that is flanked with proline at its carboxy terminus. (I) has
 CC antianemic, haemostatic and nephrotropic activities, and can be used as
 CC a stimulator of proliferation, differentiation and maturation of
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
 CC for stimulating proliferation, differentiation or growth of
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with
 CC promegakaryocytes or megakaryocytes, which results in increased platelet
 CC production. (I) with a region where amino acid residues corresponding to
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
 CC production of red blood cells, where (I) is contacted with haematopoietic
 CC stem cells or their progenitors. (I) is useful for diagnostics or
 CC therapeutics, in cell isolation strategies, and for treating patients
 CC suffering from deficiency in cell populations caused by disease,
 CC disorders or treatments related to the suppression of haematopoiesis.
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used
 in the exemplification of the present invention.
 XX
 SQ Sequence 472 AA;
 Query Match 100.0%; Score 80; DB 23; Length 472;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLROWLAARAP 15
 |||||
 DB 120 IEGPTLROWLAARAP 134
 RESULT 12
 AAW36774
 ID AAW36774 standard; peptide; 14 AA.
 AC AAW36774;
 XX
 DT 11-MAR-1998 (first entry)
 DE Thrombopoietin receptor binding peptide.
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;
 haematological disorder; thrombocytopenia; chemotherapy;
 radiation therapy; bone marrow transfusion; diagnosis;
 KW

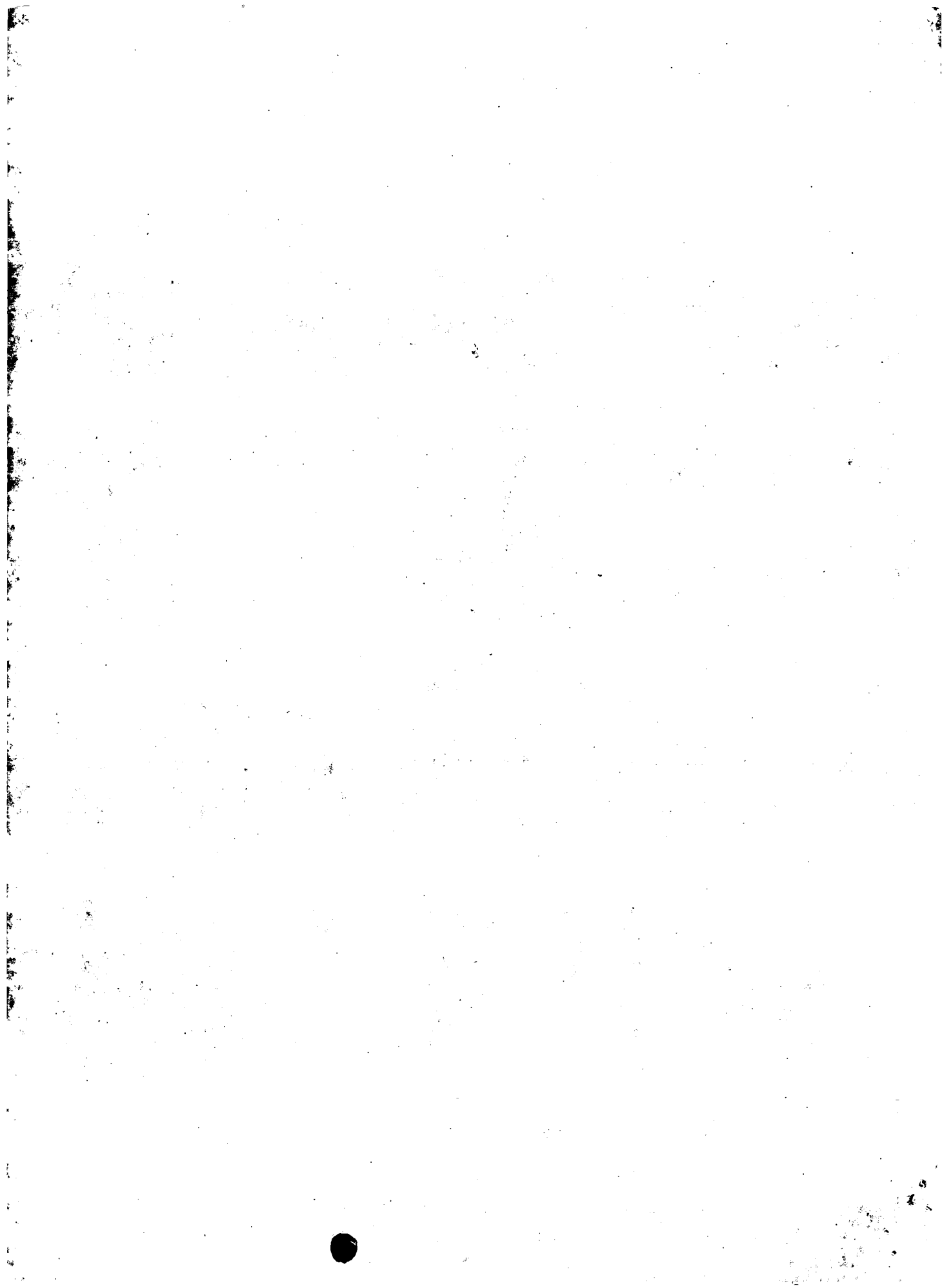
KW signal transduction; receptor activation; cell culture.
 XX Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 14
 FT /note= "NH2-Ala"
 XX
 PN WO9640750-A1.
 XX
 PD 19-DEC-1996.
 XX
 PE 07-JUN-1996; 96WO-US09623.
 PF XX
 PG 07-JUN-1995; 95US-0485301.
 PH 07-JUN-1995; 95US-0478128.
 PI (GLAX) GLAXO GROUP LTD.
 PJ Barret RW, Cwirla SE, Dower WJ, Duffin DJ, Gates CM;
 PI Johnson SS, Mattheakis LC, Schatz PJ, Wagstrom CR;
 PI Wrighton NC;
 XX WPI; 1997-052226/05.
 DR XX
 PEptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX
 PS Example 9; Page 77; 106pp; English.
 XX The present peptide, which binds the thrombopoietin receptor (TR),
 CC can be used to treat disorders which are susceptible to treatment
 CC with a thrombopoietin agonist; preferably haematological disorders
 CC and thrombocytopenia resulting from chemotherapy, radiation
 CC therapy or bone marrow transfusions. It can also be used
 CC diagnostically, e.g. to investigate the mechanism of thrombopoietin
 CC signal transduction and receptor activation, or to maintain the
 CC proliferation and growth of thrombopoietin dependent cell lines.
 XX
 SQ Sequence 14 AA;
 Query Match 91.2%; Score 73; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLROWLAARA 14
 |||||
 DB 1 IEGPTLROWLAARA 14
 RESULT 13
 AAW33030
 ID AAW33030 standard; peptide; 14 AA.
 XX
 AC AAW33030;
 XX
 DT 11-MAR-1998 (first entry)
 XX
 DE Thrombopoietin receptor binding peptide.
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;
 haematological disorder; thrombocytopenia; chemotherapy;
 radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX Synthetic.
 XX
 PN WO9640750-A1.
 XX
 PD 19-DEC-1996.
 XX
 PE 07-JUN-1996; 96WO-US09623.
 PF XX
 PG 07-JUN-1995; 95US-0485301.
 PH 07-JUN-1995; 95US-0478128.
 PI (GLAX) GLAXO GROUP LTD.
 PJ Barret RW, Cwirla SE, Dower WJ, Duffin DJ, Gates CM;
 PI Johnson SS, Mattheakis LC, Schatz PJ, Wagstrom CR;
 PI Wrighton NC;
 XX WPI; 1997-052226/05.
 DR XX
 PEptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX
 PS Example 9; Page 77; 106pp; English.
 XX The present peptide, which binds the thrombopoietin receptor (TR),
 CC can be used to treat disorders which are susceptible to treatment
 CC with a thrombopoietin agonist; preferably haematological disorders
 CC and thrombocytopenia resulting from chemotherapy, radiation
 CC therapy or bone marrow transfusions. It can also be used
 CC diagnostically, e.g. to investigate the mechanism of thrombopoietin
 CC signal transduction and receptor activation, or to maintain the
 CC proliferation and growth of thrombopoietin dependent cell lines.
 XX
 SQ Sequence 14 AA;
 Query Match 91.2%; Score 73; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLROWLAARA 14
 |||||
 DB 1 IEGPTLROWLAARA 14

XX 07-JUN-1995; 95US-0485301.
PR 07-JUN-1995; 95US-0478128.
PA (GLAX) GLAXO GROUP LTD.
XX
XX Barret RW, Cwirla SE, Dower WJ, Duffin DJ, Gates CM;
PI Johnson SS, Mattheakis LC, Schatz PJ, Wagstrom CR;
PI Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Claim 19; Page 89; 106pp; English.
PS
XX The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or
CC bone marrow transfusions. It can also be used diagnostically, e.g.
CC to investigate the mechanism of thrombopoietin signal transduction
CC and receptor activation, or to maintain the proliferation and
CC growth of thrombopoietin dependent cell lines.
XX
XX Sequence 14 AA;
SQ
Query Match 91.2%; Score 73; DB 18; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAARA 14
DB 1 IEPTLRQWLAAARA 14
RESULT 15
AAW09468
ID AAW09468 standard; protein; 14 AA.
XX
XX AAW09468;
XX
XX 10-SEP-1997 (first entry)
XX
XX Thrombopoietin receptor binding compound peptide (part of a dimer).
DE
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
KW
XX Synthetic.
XX
XX Key Location/Qualifiers
FH 14
FT Cross-links /note= "Linked to the omega Lys from AAW19534"
XX
XX WO9640189-A1.
XX
XX 19-DEC-1996.
XX
XX 05-JUN-1996; 96WO-US08998.
XX
XX 07-JUN-1995; 95US-0485301.
PR 07-JUN-1995; 95US-0472371.
PR 07-JUN-1995; 95US-0473604.
PR 07-JUN-1995; 95US-0476168.
PR 07-JUN-1995; 95US-0478128.
PR 07-JUN-1995; 95US-0484090.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Barret RW, Cwirla SE, Dower WJ, Duffin DJ, Gates CM;
PI Johnson SS, Mattheakis LC, Schatz PJ, Wagstrom CR;
PI Wrighton NC;
XX WPI; 1997-051883/05.
XX

XX 07-JUN-1995; 95US-0485301.
PR 07-JUN-1995; 95US-0478128.
PA (GLAX) GLAXO GROUP LTD.
XX
XX Barret RW, Cwirla SE, Dower WJ, Duffin DJ, Gates CM;
PI Johnson SS, Mattheakis LC, Schatz PJ, Wagstrom CR;
PI Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Claim 19; Page 89; 106pp; English.
PS
XX The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or
CC bone marrow transfusions. It can also be used diagnostically, e.g.
CC to investigate the mechanism of thrombopoietin signal transduction
CC and receptor activation, or to maintain the proliferation and
CC growth of thrombopoietin dependent cell lines.
XX
XX Sequence 14 AA;
SQ
Query Match 91.2%; Score 73; DB 18; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAARA 14
DB 1 IEPTLRQWLAAARA 14
RESULT 14
AAW33034
ID AAW33034 standard; peptide; 14 AA.
XX
XX AAW33034;
XX
XX 11-MAR-1998 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
DE
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
KW
XX Synthetic.
XX
XX Key Location/Qualifiers
FH 14
FT Cross-links /note= "terminal carboxy group linked to epsilon
FT amino group of Lys16 in AAW33035"
XX
XX WO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US09623.
XX
XX 07-JUN-1995; 95US-0485301.
PR 07-JUN-1995; 95US-0478128.
XX
XX (GLAX) GLAXO GROUP LTD.
XX

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
XX
PS Claim 30; Page 91; 106pp; English.
XX
CC The present sequence is a compound which binds to thrombopoietin (TPO)
CC receptor (TR). It is part of a dimer linked by the omega amino
CC acid to the omega amino acid in the sequence in AAW19534. The compound
CC can be used for treating patients suffering from haematological
CC disorders and thrombocytopenia resulting from chemotherapy, radiation
CC therapy or bone marrow transusions. The peptide may also be used to
CC maintain the proliferation and growth of TPO-dependent cell lines and
CC for use in biological research, for detecting TPO receptors on living
CC cells.
XX
SQ Sequence 14 AA;
SQ
SQ Every Match 91.2%; Score 73; DB 18; Length 14;
SQ Best Local Similarity 100.0%; Pred. No. 1.1e-05;
SQ Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLARA 14
DB 1 IEGPTLRQWLARA 14

Search completed: June 24, 2003, 12:47:52
Job time : 38.0769 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:45:03 ; Search time 12.4615 Seconds
(without alignments)
35.416 Million cell updates/sec

Title: US-10-006-593-2

Perfect score: 80
Sequence: 1 IEPTURQWLAARAP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Sehed: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	91.2	14	2	US-08-764-640-13
2	73	91.2	14	2	US-08-764-640-13
3	73	91.2	14	3	US-08-973-225-13
4	73	91.2	14	3	US-08-973-225-13
5	73	91.2	14	3	US-08-973-225-13
6	73	91.2	14	3	US-09-244-298A-13
7	73	91.2	14	3	US-09-244-298A-13
8	73	91.2	14	4	US-09-516-704-13
9	73	91.2	14	4	US-09-516-704-13
10	73	91.2	14	4	US-09-549-090-13
11	73	91.2	14	4	US-09-549-090-13
12	73	91.2	15	2	US-08-764-640-17
13	73	91.2	15	2	US-08-764-640-185
14	73	91.2	15	3	US-08-973-225-17
15	73	91.2	15	3	US-08-973-225-185
16	73	91.2	15	3	US-09-244-298A-17
17	73	91.2	15	4	US-09-516-704-17
18	73	91.2	15	4	US-09-516-704-185
19	73	91.2	15	4	US-09-549-090-17
20	73	91.2	15	4	US-09-549-090-185
21	73	91.2	16	2	US-08-764-640-18
22	73	91.2	16	2	US-08-764-640-194
23	73	91.2	16	2	US-08-764-640-232
24	73	91.2	16	3	US-08-973-225-18
25	73	91.2	16	3	US-08-973-225-194
26	73	91.2	16	3	US-08-973-225-220
27	73	91.2	16	3	US-09-244-298A-18

28	73	91.2	16	3	US-09-244-298A-194
29	73	91.2	16	3	US-09-244-298A-232
30	73	91.2	16	4	US-09-516-704-18
31	73	91.2	16	4	US-09-516-704-194
32	73	91.2	16	4	US-09-516-704-232
33	73	91.2	16	4	US-09-549-090-18
34	73	91.2	16	4	US-09-549-090-194
35	73	91.2	16	4	US-09-549-090-220
36	69	86.2	14	2	US-08-764-640-195
37	69	86.2	14	2	US-08-764-640-199
38	69	86.2	14	3	US-08-973-225-195
39	69	86.2	14	3	US-08-973-225-199
40	69	86.2	14	3	US-09-244-298A-195
41	69	86.2	14	3	US-09-244-298A-199
42	69	86.2	14	4	US-09-516-704-195
43	69	86.2	14	4	US-09-516-704-199
44	69	86.2	14	4	US-09-549-090-195
45	69	86.2	14	4	US-09-549-090-199

ALIGNMENTS

RESULT 1
US-08-764-640-13
; Sequence 13, Application US/08764640
; Patent No. 5869451
; Patent No. 5869451 5837683
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Gates, Christian
; APPLICANT: Schatz, Peter J.
; APPLICANT: Balasubramanian, Palaniappan
; APPLICANT: Wagstrom, Christopher R.
; APPLICANT: Hendren, Richard W.
; APPLICANT: Deprence, Randolph B.
; APPLICANT: Podduturi, Surekha
; APPLICANT: Yin, Qun
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,640
; FILING DATE: 11-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-764-640-13

Query Match 91.2%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e-05; 0; Mismatches 0; Gaps 0;
Matches 14; Conservative 0; Indels 0;

QY 1 IEPTLRQWLAARA 14
DB 1 IEPTLRQWLAARA 14

RESULT 2

US-08-764-640-193

; Sequence 193, Application US/08764640

; Patent No. 5869451

; Patent No. 5869451 5837683

; GENERAL INFORMATION:

; APPLICANT: Dower, William J.

; APPLICANT: Barrett, Ronald W.

; APPLICANT: Cwiria, Steven E.

; APPLICANT: Gates, Christian

; APPLICANT: Schatz, Peter J.

; APPLICANT: Balasubramanian, Palaniappan

; APPLICANT: Wagstrom, Christopher R.

; APPLICANT: Hendren, Richard W.

; APPLICANT: Deprence, Randolph B.

; APPLICANT: Podduturi, Surekha

; APPLICANT: Yin, Qun

; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A

; TITLE OF INVENTION: RECEPTOR

; NUMBER OF SEQUENCES: 244

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Glaxo Wellcome

; STREET: Five Moore Drive, P.O. Box 13398

; CITY: Research Triangle Park

; STATE: NC

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/764,640

; FILING DATE: 11-DEC-1996

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Hrubiec, Robert T.

; REGISTRATION NUMBER: 36,392

; REFERENCE/DOCKET NUMBER: PK3281

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-248-1000

; INFORMATION FOR SEQ ID NO: 193:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 14 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-764-640-193

Query Match 91.2%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
DB 1 IEPTLRQWLAARA 14

RESULT 3

US-08-973-225-13

; Sequence 13, Application US/08973225A

; Patent No. 6083913

; GENERAL INFORMATION:

; APPLICANT: Dower, William J.

; APPLICANT: Barrett, Ronald W.

; APPLICANT: Cwiria, Steven E.

; APPLICANT: Duffin, David J.

; APPLICANT: Gates, Christian

; APPLICANT: Haselden, Sherril S.

; APPLICANT: Mattheakis, Larry C.

; APPLICANT: Schatz, Peter J.

; APPLICANT: Wagstrom, Christopher R.

; APPLICANT: Wrighton, Nicholas C.

; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A

; TITLE OF INVENTION: THROMBOPOIETIN RECEPTOR

; NUMBER OF SEQUENCES: 232

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Glaxo Wellcome

; STREET: Five Moore Drive, P.O. Box 13398

; CITY: Research Triangle Park

; STATE: NC

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/973,225A

; FILING DATE: 04-Dec-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Hrubiec, Robert T.

; REGISTRATION NUMBER: 36,392

; REFERENCE/DOCKET NUMBER: PK3065USW

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-248-1000

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 14 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 13:

; US-08-973-225-13

Query Match 91.2%; Score 73; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
DB 1 IEPTLRQWLAARA 14

RESULT 4

US-08-973-225-193

; Sequence 193, Application US/08973225A

; Patent No. 6083913

; GENERAL INFORMATION:

; APPLICANT: Dower, William J.

; APPLICANT: Barrett, Ronald W.

; APPLICANT: Cwiria, Steven E.

; APPLICANT: Duffin, David J.

; APPLICANT: Gates, Christian

; APPLICANT: Haselden, Sherril S.

; APPLICANT: Mattheakis, Larry C.

; APPLICANT: Schatz, Peter J.

; APPLICANT: Wagstrom, Christopher R.

; APPLICANT: Wrighton, Nicholas C.

; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A

; TITLE OF INVENTION: THROMBOPOIETIN RECEPTOR

; NUMBER OF SEQUENCES: 232

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,225A
; FILING DATE: 04-Dec-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3065USW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-08-973-225-193

Query Match 91.2%; Score 73; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAAR 14
Db 1 IEGPTLRQWLAAAR 14

RESULT 5
US-09-244-298A-13
; Sequence 13, Application US/09244298A
; Patent No. 6121238
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Gates, Christian
; APPLICANT: Schatz, Peter J.
; APPLICANT: Balasubramanian, Palaniappan
; APPLICANT: Wagstrom, Christopher R.
; APPLICANT: Hendren, Richard W.
; APPLICANT: Deprince, Randolph B.
; APPLICANT: Podduturi, Surekha
; APPLICANT: Yin, Qun
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/244,298A
; FILING DATE: 11-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

; FILING DATE: 11-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-244-298A-13

Query Match 91.2%; Score 73; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAAR 14
Db 1 IEGPTLRQWLAAAR 14

RESULT 6
US-09-244-298A-193
; Sequence 193, Application US/09244298A
; Patent No. 6121238
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Gates, Christian
; APPLICANT: Schatz, Peter J.
; APPLICANT: Balasubramanian, Palaniappan
; APPLICANT: Wagstrom, Christopher R.
; APPLICANT: Hendren, Richard W.
; APPLICANT: Deprince, Randolph B.
; APPLICANT: Podduturi, Surekha
; APPLICANT: Yin, Qun
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/244,298A
; FILING DATE: 11-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

MOLECULE TYPE: peptide
US-09-244-298A-193

Query Match 91.2%; Score 73; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e-05; 0; Mismatches 0; Gaps 0;
Matches 14; Conservative 0; Indels 0;

QY 1 IEPTLRQWLAARA 14
Db 1 IEPTLRQWLAARA 14

RESULT 7
US-09-516-704-13
; Sequence 13, Application US/09516704
; Patent No. 6251864
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; Barrett, Ronald W.
; Cwirla, Steven E.
; Gates, Christian
; Schatz, Peter J.
; Balasubramanian, Palaniappan
; Wagstrom, Christopher R.
; Hendren, Richard W.
; DePrince, Randolph B.
; Poduturi, Surekha
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/516.704
; FILING DATE: 01-Mar-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-516-704-13

Query Match 91.2%; Score 73; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
Db 1 IEPTLRQWLAARA 14

RESULT 8
US-09-516-704-193
; Sequence 13, Application US/09549090
; Patent No. 6465430
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; Barrett, Ronald W.
; Cwirla, Steven E.
; Duffin, David J.
; Gates, Christian
; Haselden, Sherril S.
; Matheakis, Larry C.
; Schatz, Peter J.
; Wagstrom, Christopher R.
; Wrighton, Nicholas C.
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A

THROMBOPOIETIN RECEPTOR
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSEE: Glaxo Wellcome
STREET: Five Moore Drive, P.O. Box 13398
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 13-Apr-2000
APPLICATION NUMBER: US/09/549,090
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/973,225
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hrubiec, Robert T.
REGISTRATION NUMBER: 36,392
REFERENCE/DOCKET NUMBER: PK3065USW
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-248-1000
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-549-090-13
Query Match 91.2%; Score 73; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. NO. 1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAARA 14
Db 1 IEPTLRQWLAAARA 14
RESULT 10
09-549-090-193
Sequence 193, Application US/09549090
Patent No. 6465430
GENERAL INFORMATION:
APPLICANT: Dower, William J.
Barrett, Ronald W.
Cwiria, Steven E.
Duffin, David J.
Gates, Christian
Haselden, Sherril S.
Mattheakis, Larry C.
Schatz, Peter J.
Wagstrom, Christopher R.
Wrighton, Nicholas C.
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
THROMBOPOIETIN RECEPTOR
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSEE: Glaxo Wellcome
STREET: Five Moore Drive, P.O. Box 13398
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/549,090
FILING DATE: 13-Apr-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/973,225
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hrubiec, Robert T.
REGISTRATION NUMBER: 36,392
REFERENCE/DOCKET NUMBER: PK3065USW
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-248-1000
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-549-090-193
Query Match 91.2%; Score 73; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. NO. 1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAARA 14
Db 1 IEPTLRQWLAAARA 14
RESULT 11
US-08-764-640-17
Sequence 17, Application US/08764640
Patent No. 5869451
Patent No. 5869451 5837683
GENERAL INFORMATION:
APPLICANT: Dower, William J.
APPLICANT: Barrett, Ronald W.
APPLICANT: Cwiria, Steven E.
APPLICANT: Gates, Christian
APPLICANT: Schatz, Peter J.
APPLICANT: Balasubramanian, Palaniappan
APPLICANT: Wagstrom, Christopher R.
APPLICANT: Hendren, Richard W.
APPLICANT: Deprience, Randolph B.
APPLICANT: Podduturi, Surekha
APPLICANT: Yin, Qun
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
TITLE OF INVENTION: RECEPTOR
NUMBER OF SEQUENCES: 244
CORRESPONDENCE ADDRESS:
ADDRESSEE: Glaxo Wellcome
STREET: Five Moore Drive, P.O. Box 13398
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,640
FILING DATE: 11-DEC-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Hrubiec, Robert T.
REGISTRATION NUMBER: 36,392
REFERENCE/DOCKET NUMBER: PK3281

TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-248-1000
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-764-640-17

Query Match 91.2%; Score 73; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
| | | | | | | | | | | | | | | |
DB 1 IEPTLRQWLAARA 14
| | | | | | | | | | | | | | | |

RESULT 12
US-08-764-640-185

Sequence 185, Application US/08764640
Patent No. 5869451
Patent No. 5869451 5837683

GENERAL INFORMATION:

APPLICANT: Dower, William J.
APPLICANT: Barrett, Ronald W.
APPLICANT: Cwiria, Steven E.
APPLICANT: Gates, Christian
APPLICANT: Schatz, Peter J.
APPLICANT: Balasubramanian, Palaniappan
APPLICANT: Wagstrom, Christopher R.
APPLICANT: Hendren, Richard B.
APPLICANT: Deprence, Randolph B.
APPLICANT: Poduturi, Surekha
APPLICANT: Yin, Qun

TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
RECEPTOR
NUMBER OF SEQUENCES: 244

CORRESPONDENCE ADDRESS:
ADDRESSEE: Glaxo Wellcome
STREET: Five Moore Drive, P.O. Box 13398
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,640
FILING DATE: 11-DEC-1996
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Hrubiec, Robert T.
REGISTRATION NUMBER: 36,392
REFERENCE/DOCKET NUMBER: PK3281
TELEPHONE: 919-248-1000
INFORMATION FOR SEQ ID NO: 185:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-764-640-185

Query Match 91.2%; Score 73; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAARA 14
| | | | | | | | | | | | | | | |
DB 2 IEPTLRQWLAARA 15
| | | | | | | | | | | | | | | |

RESULT 13
US-08-973-225-17

Sequence 17, Application US/08973225A
Patent No. 6083913

GENERAL INFORMATION:

APPLICANT: Dower, William J.
APPLICANT: Barrett, Ronald W.
APPLICANT: Cwiria, Steven E.
APPLICANT: Duffin, David J.
APPLICANT: Gates, Christian
APPLICANT: Haselden, Sherrill S.
APPLICANT: Mattheakis, Larry C.
APPLICANT: Schatz, Peter J.
APPLICANT: Wagstrom, Christopher R.
APPLICANT: Wrighton, Nicholas C.

TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
THROMBOPOIETIN RECEPTOR

NUMBER OF SEQUENCES: 232

CORRESPONDENCE ADDRESS:

ADDRESSEE: Glaxo Wellcome
STREET: Five Moore Drive, P.O. Box 13398
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,225A
FILING DATE: 04-DEC-1997

ATTORNEY/AGENT INFORMATION:
NAME: Hrubiec, Robert T.
REGISTRATION NUMBER: 36,392
REFERENCE/DOCKET NUMBER: PK3065USW
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-248-1000
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-08-973-225-17

Query Match 91.2%; Score 73; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
| | | | | | | | | | | | | | | |
DB 1 IEPTLRQWLAARA 14
| | | | | | | | | | | | | | | |

RESULT 14
US-08-973-225-185

Sequence 185, Application US/08973225A
Patent No. 6083913

GENERAL INFORMATION:

APPLICANT: Dower, William J.
APPLICANT: Barrett, Ronald W.
APPLICANT: Cwiria, Steven E.

;; Duffin, David J.
;; Gates, Christian
;; Haselden, Sherril S.
;; Mattheakis, Larry C.
;; Schatz, Peter J.
;; Wagstrom, Christopher R.
;; Wrighton, Nicholas C.
;; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;; THROMBOPOIETIN RECEPTOR

;; NUMBER OF SEQUENCES: 232
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Glaxo Wellcome
;; STREET: Five Moore Drive, P.O. Box 13398
;; CITY: Research Triangle Park

;; STATE: NC

;; COUNTRY: USA

;; ZIP: 27709

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent In Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/973,225A

;; FILING DATE: 04-Dec-1997

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Hrubiec, Robert T.

;; REGISTRATION NUMBER: 36,392

;; REFERENCE/DOCKET NUMBER: PK3065USW

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 919-248-1000

;; INFORMATION FOR SEQ ID NO: 185:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 15 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: <Unknown>

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; SEQUENCE DESCRIPTION: SEQ ID NO: 185:

US-08-973-225-185

Query Match 91.2%; Score 73; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGLTLRQWLAAARA 14
| | | | | | | | | | | | | | | |
DB 2 IEGLTLRQWLAAARA 15

RESULT 15

US-09-244-298A-17

;; Sequence 17, Application US/09244298A

;; Patent No. 6121238

;; GENERAL INFORMATION:

;; APPLICANT: Dower, William J.

;; APPLICANT: Barrett, Ronald W.

;; APPLICANT: Cwiria, Steven E.

;; APPLICANT: Gates, Christian

;; APPLICANT: Schatz, Peter J.

;; APPLICANT: Balasubramanian, Palaniappan

;; APPLICANT: Wagstrom, Christopher R.

;; APPLICANT: Hendren, Richard W.

;; APPLICANT: Deprence, Randolph B.

;; APPLICANT: Podduturi, Surekha

;; APPLICANT: Yin, Qun

;; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A

;; RECEPTOR

;; NUMBER OF SEQUENCES: 244

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Glaxo Wellcome

;; STREET: Five Moore Drive, P.O. Box 13398

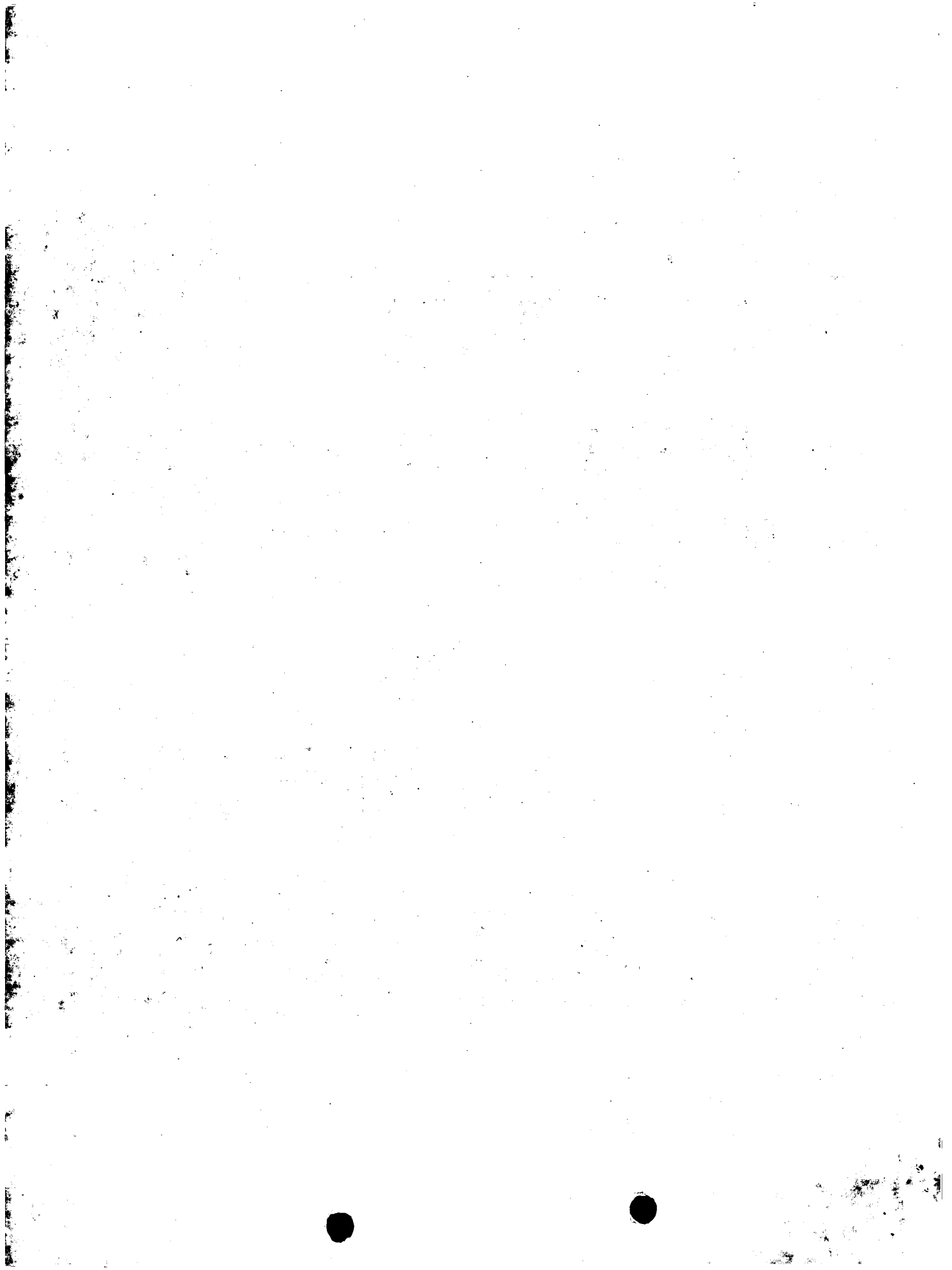
;; CITY: Research Triangle Park

;; STATE: NC
;; COUNTRY: USA
;; ZIP: 27709
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/244,298A
;; FILING DATE: 11-Dec-1996
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hrubiec, Robert T.
;; REGISTRATION NUMBER: 36,392
;; REFERENCE/DOCKET NUMBER: PK3281
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-248-1000
;; INFORMATION FOR SEQ ID NO: 17:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-09-244-298A-17

Query Match 91.2%; Score 73; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGLTLRQWLAAARA 14
| | | | | | | | | | | | | | | |
DB 1 IEGLTLRQWLAAARA 14

Search completed: June 24, 2003, 12:53:25
Job time : 13.4615 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:50:04 ; Search time 20.7692 Seconds
(without alignments)
78.149 Million cell updates/sec

Title: US-10-006-593-2

Perfect score: 80
Sequence: 1 IEGLTLRWLAARAP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Shed: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	15	9 US-10-006-593-2	Sequence 2, Appli
2	80	100.0	18	9 US-10-006-593-31	Sequence 31, Appl
3	80	100.0	18	9 US-10-006-593-35	Sequence 35, Appl
4	80	100.0	18	9 US-10-006-593-37	Sequence 37, Appl
5	80	100.0	18	9 US-10-006-593-39	Sequence 39, Appl
6	80	100.0	18	9 US-10-006-593-41	Sequence 41, Appl
7	80	100.0	18	9 US-10-006-593-43	Sequence 43, Appl
8	80	100.0	18	9 US-10-006-593-45	Sequence 45, Appl
9	80	100.0	18	9 US-10-006-593-49	Sequence 49, Appl
10	80	100.0	18	9 US-10-006-593-66	Sequence 66, Appl
11	80	100.0	472	9 US-10-006-593-67	Sequence 67, Appl
12	73	91.2	14	9 US-10-006-593-1	Sequence 1, Appli
13	73	91.2	14	9 US-10-304-160-31	Sequence 31, Appl
14	73	91.2	18	9 US-10-006-593-25	Sequence 25, Appl
15	73	91.2	18	9 US-10-006-593-27	Sequence 27, Appl
16	73	91.2	18	9 US-10-006-593-29	Sequence 29, Appl
17	73	91.2	18	9 US-10-006-593-33	Sequence 33, Appl
18	73	91.2	18	9 US-10-006-593-47	Sequence 47, Appl
19	73	91.2	18	9 US-10-006-593-61	Sequence 61, Appl

20	73	91.2	28	9 US-10-006-593-56	Sequence 56, Appl
21	48.5	60.6	607	9 US-10-014-717-9	Sequence 9, Appli
22	47	58.8	8	9 US-10-006-593-112	Sequence 112, App
23	47	58.8	18	9 US-10-006-593-113	Sequence 113, App
24	46	57.5	1256	9 US-10-047-542-90	Sequence 90, Appl
25	45	56.2	1252	9 US-10-047-542-89	Sequence 89, Appl
26	43	53.8	660	10 US-09-815-242-5470	Sequence 5470, Ap
27	43	53.8	664	10 US-09-815-242-12179	Sequence 12179, A
28	41	51.2	211	10 US-09-925-301-1056	Sequence 1056, Ap
29	41	51.2	308	9 US-09-738-626-5723	Sequence 5723, Ap
30	40.5	50.6	125	9 US-09-738-626-4661	Sequence 4661, Ap
31	40	50.0	40	10 US-09-864-761-44217	Sequence 44217, A
32	40	50.0	716	9 US-09-738-626-5884	Sequence 5884, Ap
33	40	50.0	1404	10 US-09-862-027-24	Sequence 24, Appl
34	40	50.0	1712	9 US-09-961-403-9	Sequence 9, Appli
35	39	48.8	267	9 US-09-813-453A-15	Sequence 15, Appl
36	38.5	48.1	156	9 US-10-074-475-237	Sequence 237, App
37	38	47.5	78	9 US-10-001-857-172	Sequence 172, App
38	38	47.5	443	12 US-10-083-452-5	Sequence 5, Appli
39	38	47.5	905	9 US-09-738-626-4994	Sequence 4994, Ap
40	37	46.2	122	9 US-10-114-893-180	Sequence 180, App
41	37	46.2	217	10 US-09-799-777-26	Sequence 26, Appl
42	37	46.2	304	9 US-10-013-477-11	Sequence 11, Appl
43	37	46.2	318	9 US-10-001-254-18	Sequence 18, Appl
44	37	46.2	319	10 US-09-925-302-758	Sequence 758, App
45	37	46.2	366	9 US-10-106-698-4626	Sequence 4626, Ap

ALIGNMENTS

RESULT 1
US-10-006-593-2
; Sequence 2, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
US-10-006-593-2
Query Match 100.0%; Score 80; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGLTLRWLAARAP 15
| | | | | | | | | | | | | | | |
Db 1 IEGLTLRWLAARAP 15
RESULT 2
US-10-006-593-31
; Sequence 31, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:

```
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 31
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
; US-10-006-593-31
Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAAP 15
DB 3 IEPTLRQWLAAAP 17
RESULT 3
US-10-006-593-35
; Sequence 35, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 35
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
; US-10-006-593-35
Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAAP 15
DB 3 IEPTLRQWLAAAP 17
RESULT 4
US-10-006-593-37
; Sequence 37, Application US/10006593
; Publication No. US20030049683A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 37
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
; US-10-006-593-37
Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAAP 15
DB 3 IEPTLRQWLAAAP 17
RESULT 5
US-10-006-593-39
; Sequence 39, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 39
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
; US-10-006-593-39
Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEPTLRQWLAAAP 15
DB 3 IEPTLRQWLAAAP 17
RESULT 6
US-10-006-593-41
; Sequence 41, Application US/10006593
```

; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
US-10-006-593-41

Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAAP 15
Db 3 IEGPTLRQWLAAAP 17

RESULT 7

US-10-006-593-43
; Sequence 43, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
US-10-006-593-43

Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAAP 15
Db 3 IEGPTLRQWLAAAP 17

RESULT 8

US-10-006-593-45

; Sequence 45, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
US-10-006-593-45

Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAAP 15
Db 3 IEGPTLRQWLAAAP 17

RESULT 9

US-10-006-593-49
; Sequence 49, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
US-10-006-593-49

Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAAP 15
Db 3 IEGPTLRQWLAAAP 17

RESULT 10

US-10-006-593-66
; Sequence 66, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 66
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic peptide with flanking sequence
US-10-006-593-66

Query Match 100.0%; Score 80; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARAP 15
|||
DB 3 IEGPTLRQWLAAARAP 17
|||

RESULT 11

US-10-006-593-67
; Sequence 67, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 67
; LENGTH: 472
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Humanized antibody heavy chain
US-10-006-593-67

Query Match 100.0%; Score 80; DB 9; Length 472;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARAP 15
|||
DB 120 IEGPTLRQWLAAARAP 134
|||

RESULT 12
US-10-006-593-1
; Sequence 1, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 14
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO/mimetic peptide
US-10-006-593-1

Query Match 91.2%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
|||
DB 1 IEGPTLRQWLAAARA 14
|||

RESULT 13

US-10-304-160-31
; Sequence 31, Application US/10304160
; Publication No. US20030099619A1
; GENERAL INFORMATION:
; APPLICANT: WICKHAM, THOMAS J
; APPLICANT: KOVESDI, IMRE
; APPLICANT: ROELVINK, PETRUS W
; APPLICANT: EINFELD, DAVID
; APPLICANT: BROUGH, DOUGLAS E
; APPLICANT: LIZONOVA, ALENA
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR TARGETING AN ADENOVIRAL VECTOR
; FILE REFERENCE: 220148
; CURRENT APPLICATION NUMBER: US/10/304,160
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: PCT/US01/17391
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 09/631,191
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: US 60/208451
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 3.1
; SEQ ID NO 31
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-304-160-31

Query Match 91.2%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
|||
DB 1 IEGPTLRQWLAAARA 14
|||

Db 3 IEGPTLRQWLAARA 16

Search completed: June 24, 2003, 13:10:45
Job time : 20.7692 secs

RESULT 14
US-10-006-593-25
; Sequence 25, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
US-10-006-593-25

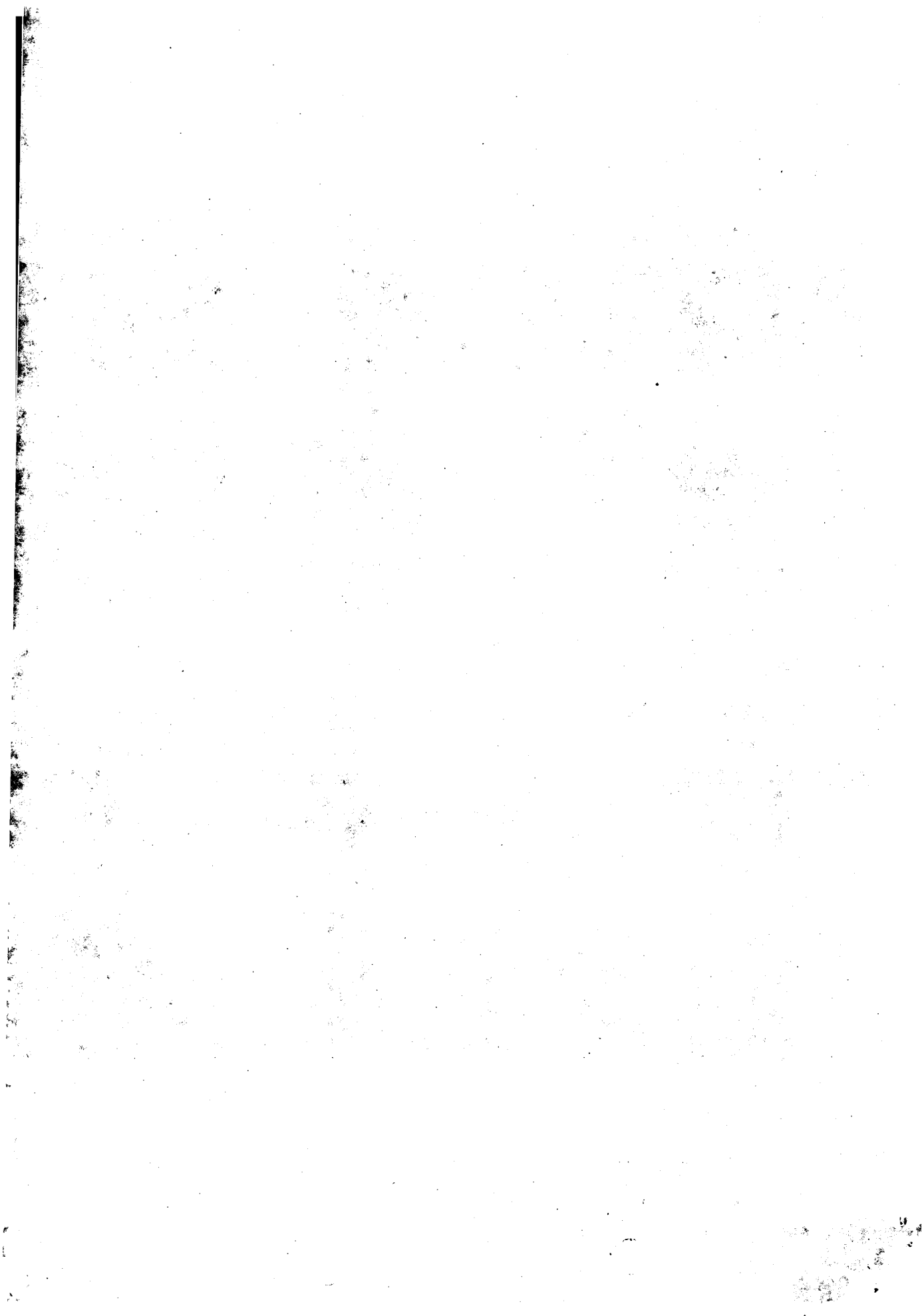
Query Match 91.2%; Score 73; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | | |
Db 3 IEGPTLRQWLAARA 16

RESULT 15
US-10-006-593-27
; Sequence 27, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 18
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: TPO mimetic with flanking amino acids
US-10-006-593-27

Query Match 91.2%; Score 73; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | | |



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:42:13 ; Search time 16.1538 Seconds
(without alignments)
89.268 Million cell updates/sec

Title: US-10-006-593-2

Perfect score: 80
Sequence: 1 IEQPTLRQWLARAP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Shed: 283224 seqs, 95134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:*

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.5	61.9	333	2 A36925	transcription acti
2	47	58.8	296	2 AG0147	probable membrane
3	46	57.5	306	2 D70601	UTP-glucose-1-phos
4	46	57.5	400	2 C87021	serine-threonine p
5	44.5	55.6	436	2 JC4742	transposase - Cory
6	44	55.0	200	2 T23485	hypothetical prote
7	44	55.0	207	2 T37464	probable glutathio
8	44	55.0	536	1 SYECEB	2,3-dihydroxybenzo
9	44	55.0	536	2 A99708	2,3-dihydroxybenzo
10	44	55.0	536	2 E85558	serine/threonine p
11	44	55.0	664	2 H83962	cis/trans isomerases
12	44	55.0	762	2 H83415	thiamine-phosphate
13	43	53.8	204	2 T44257	probable dimethyla
14	43	53.8	285	2 G71337	protein kinase [im
15	43	53.8	664	2 G89894	conserved hypotet
16	43	53.8	683	2 B71325	DNA-binding protei
17	43	53.8	825	2 JC4163	membrane protein,
18	42	52.5	352	2 G83636	hypothetical prote
19	42	52.5	438	2 G87337	pol polyprotein -
20	42	52.5	473	2 E84853	peptide synthetase
21	42	52.5	1019	2 T11560	probable phosphopa
22	42	52.5	2617	2 AE2136	probable phosphopa
23	41	51.2	195	2 F91171	hypothetical 21.8K
24	41	51.2	195	2 F86017	ABC transporter, A
25	41	51.2	195	2 S47694	UTP-glucose-1-phos
26	41	51.2	249	2 E8575	glyceraldehyde-3-p
27	41	51.2	306	2 T45453	glyceraldehyde-3-p
28	41	51.2	326	2 C24430	
29	41	51.2	336	1 DEP2G	

30 41 51.2 337 2 A35080 glyceraldehyde-3-p
31 41 51.2 338 1 DEIS3C glyceraldehyde-3-p
32 41 51.2 338 2 JQ1287 glyceraldehyde-3-p
33 41 51.2 399 1 B70936 probable serine/th
34 41 51.2 401 2 G87552 conserved hypotet
35 41 51.2 719 2 B95325 conserved hypotet
36 41 51.2 750 2 A97501 topoisomerase iv c
37 41 51.2 750 2 AE2719 topoisomerase iv c
38 40 50.0 239 2 S25204 smx protein - Str
39 40 50.0 321 2 C70653 probable prephenat
40 40 50.0 463 2 S27491 hypothetical prote
41 40 50.0 518 2 AD2315 hypothetical prote
42 40 50.0 530 2 AB1958 probable permease
43 40 50.0 531 2 E81015 ABC transporter, p
44 40 50.0 648 1 H69878 probable protein k
45 40 50.0 656 2 S30483 pol polyprotein -

ALIGNMENTS

RESULT 1

A36925

transcription activator LysR-type Cbbr - Xanthobacter flavus

C;Species: Xanthobacter flavus

C;Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text_change 24-Sep-1999

C;Accession: A36925; S13578; S35408

R;van den Bergh, E.R.E.; Dijkhuizen, L.; Meijer, W.G.

J. Bacteriol. 175, 6097-6104, 1993

A;Title: Cbbr, a LysR-type transcriptional activator, is required for expression of the

A;Reference number: A36925; MUID:94012468; PMID:8407781

A;Accession: A36925

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-333 <VAN>

A;Cross-references: EMBL:Z22705; NID:g297851; PIDN:CAA80406.1; PID:g581832

R;Meijer, W.G.; Arnberg, A.C.; Enequist, H.G.; Terpstra, P.; Lidstrom, M.E.; Dijkhuizen

Mol. Gen. Genet. 225, 320-330, 1991

A;Title: Identification and organization of carbon dioxide fixation genes in Xanthobact

A;Reference number: S13573; MUID:91172133; PMID:1900916

A;Accession: S13578

A;Molecule type: DNA

A;Residues: 1-150 <MEI>

A;Cross-references: EMBL:X17252

C;Genetics:

A;Gene: cbbR

A;Start codon: GTG

C;Superfamily: transcription activator LysR-type

C;Keywords: DNA binding; transcription regulation

Query Match 61.9%; Score 49.5; DB 2; Length 333;

Best Local Similarity 66.7%; Pred. No. 1.5;

Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 IEQ-PTLRQWLARA 14

264 VEGLPVVRQWLAVRA 278

RESULT 2

AG0147

probable membrane protein YPO1203 [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002

C;Accession: AG0147

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB00001; MUID:21470413; PMID:11586360

A;Accession: AG0147

A;Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-296 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC90042.1; PID:G15979263; GSPDB:GN00175
 C:Genetics:
 C:Superfamily: hypothetical protein ydeD

Query Match 58.8%; Score 47; DB 2; Length 296;
 Best Local Similarity 81.8%; Pred. No. 3.5;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 PTLROWLAARA 14
 DB 66 PTLROWAASA 76

RESULT 3
 D70601
 UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - Mycobacterium
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 03-Jun-2002
 C:Accession: D70601
 Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: D70601
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-306 <COL>
 A:Cross-references: GB:Z94752; GB:AL123456; NID:G3261731; PIDN:CAB08153.1; PID:G2052127
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: galU
 C:Superfamily: Escherichia coli UTP-glucose-1-phosphate uridylyltransferase
 C:Keywords: nucleotidyltransferase

Query Match 57.5%; Score 46; DB 2; Length 306;
 Best Local Similarity 72.7%; Pred. No. 5.3;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 GPTLROWLAAR 13
 DB 290 GPTLRLRLVAR 300

RESULT 4
 C87021
 serine-threonine protein kinase (imported) - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 10-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-May-2002
 C:Accession: C87021
 R; Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hamlin, N.; Devlin, K.; Feltwell, T.; Duthoy, S.; Fraser, A.; Hamlin, N.; Holtroyd, S.; Nature 409, 1007-1011, 2001
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squares, R.; Nature 409, 1007-1011, 2001
 A:Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: C87021; A86509; MUID:21128732; PMID:11234002
 A:Accession: C87021
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-400 <SPO>
 A:Cross-references: GB:AL450380; NID:G13092968; PIDN:CAC31278.1; GSPDB:GN00147
 A:Gene: ML0897
 C:Superfamily: Mycobacterium tuberculosis probable serine/threonine-specific protein kinase

Query Match 57.5%; Score 46; DB 2; Length 400;
 Best Local Similarity 66.7%; Pred. No. 7;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 IEPTLROWLAARAP 15
 DB 86 IEPTLRELLAARGP 100

RESULT 5
 J24742
 transposase - Corynebacterium glutamicum
 C:Species: Corynebacterium glutamicum
 C:Date: 0-May-1996 #sequence_revision 16-Aug-1996 #text_change 17-Mar-1999
 C:Accession: J24742
 R; Corbett, A.; Fieshaber, A.; Castro, J.M.; Martin, J.F.
 Gene 170, 91-94, 1996
 A:Title: Cloning and characterization of an IS-like element present in the genome of Br.
 A:Reference number: J24742; MUID:96200862; PMID:8621097
 A:Accession: J24742
 A:Molecule type: DNA
 A:Residues: 1-436 <COR>
 A:Cross-references: EMBL:Z66534
 A:Experimental source: ATCC 13669
 A:Note: The authors translated the initiation codon TGT for residue 1 as Val
 A:Note: The authors translated the codon ATT for residue 125 as Tyr
 A:Note: The source is designated as Brevibacterium lactofermentum
 C:Genetics:
 A:Gene: GTG
 F:388-415/Domain: DNA binding #status predicted <DNA>
 F:405-415/Region: helix-turn-helix

Query Match 55.6%; Score 44.5; DB 2; Length 436;
 Best Local Similarity 55.6%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 4; Indels 3; Gaps 1;
 QY 1 IEPTLROWLAARAP 15
 DB 206 VEGRSADALRTWLAARTP 223

RESULT 6
 T23485
 hypothetical protein K08F4.11 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T23485
 R; Hembry, C.
 submitted to the EMBL Data Library, January 1996
 A:Reference number: Z19746
 A:Accession: T23485
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-200 <WIL>
 A:Cross-references: EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN000022; CESP:K08F4.11
 A:Experimental source: clone K08F4
 C:Genetics:
 A:Gene: CESP:K08F4.11
 A:Map position: 4
 A:Introns: 45/1; 76/1; 111/3
 C:Superfamily: glutathione transferase

Query Match 55.0%; Score 44; DB 2; Length 200;
 Best Local Similarity 61.5%; Pred. No. 7.4;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 IEPTLROWLAAR 13
 DB 183 IETPLKEWLAAR 195

RESULT 7
 T37464
 probable glutathione transferase (EC 2.5.1.18) GST3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T37464
R;Tawe, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehrsen, K.
submitted to the EMBL Data Library, June 1997
A;Description: Paraquat mediates differential gene expression in *C. elegans*.
A;Reference number: Z21702
A;Accession: T37464
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-207 <TAW>
A;Cross-references: EMBL:AF010241; PIDN:AAB65419.1
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: GST3
C;Superfamily: glutathione transferase
C;Keywords: transferase

Query Match 55.0%; Score 44; DB 2; Length 207;
Best Local Similarity 61.5%; Pred. No. 7.6;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAAR 13
:::|||||
Db 190 IETPKLEWLAKR 202
:::|||||

RESULT 8

SYCEB
2,3-dihydroxybenzoate-[carrier protein] ligase (EC 6.2.1.-) entE - Escherichia coli (str
N;Alternate names: 2,3-dihydroxybenzoate-AMP ligase [mismomer]; dihydroxybenzoic acid-ac
C;Species: Escherichia coli
C;Date: 31-Dec-1989 #sequence_revision 21-Nov-1997 #text_change 01-Mar-2002
C;Accession: H64792; A48308; A32047; I41058; S08076
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
-A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64792
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-536 <BLAT>
A;Cross-references: GB:AE000165; GB:U00096; NID:G1786808; PIDN:AAC73695.1; PID:G1786810;
A;Experimental source: strain K-12, substrain MG1655
R;Staab, J.F.; Elkins, M.F.; Barhart, C.F.
FEMS Microbiol. Lett. 59, 15-19, 1989

A;Title: Nucleotide sequence of the Escherichia coli entE gene.
A;Reference number: A48308; MUID:89290355; PMID:2525505
A;Note: in Medline 89290355 this citation is erroneously given as volume 50 rather than
A;Accession: A48308
A;Molecule type: DNA
A;Residues: 1-368, 'ECRRKSTAAAR', 379-536 <STA>
A;Cross-references: GB:M27490; EMBL:X15058; NID:G41345; PIDN:CAA33158.1; PID:G41346
R;Liu, J.; Duncan, K.; Walsh, C.T.
J. Bacteriol. 171, 791-798, 1989

A;Title: Nucleotide sequence of a cluster of Escherichia coli enterobactin biosynthesis
A;Reference number: A91904; MUID:89123155; PMID:2521622
A;Accession: A32047
A;Molecule type: DNA
A;Residues: 333-536 <LIU>
A;Cross-references: GB:M24148; NID:G304949; PIDN:AAAL6101.1; PID:G450380
C;Comment: The enzymatic steps in the condensation of L-serine and 2,3-dihydroxybenzoic
ty is based on its recognized homology with 4-coumarate-CoA ligase and by analogy with
C;Comment: The formation of 2,3-dihydroxybenzoyl-AMP has been observed. The rapid reacti
carrier protein) to release AMP, has also been observed.
C;Genetics:
A;Gene: entE
A;Map position: 14 min
C;Function:
A;Description: catalyzes the formation of 2,3-dihydroxybenzoyl-(carrier protein), AMP an
A;Pathway: enterobactin biosynthesis
A;Note: this is one component of a membrane-bound multienzyme complex that catalyzes the
for transport into the cell

C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
C;Keywords: acid-thiol ligase; enterobactin biosynthesis; membrane-associated complex
F;69-526/Domain: acetate-CoA ligase homology <ACL>

Query Match 55.0%; Score 44; DB 1; Length 536;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEQPTLRQWLAAAR 14
:::|||||
Db 521 VDKKQLRQWLASRA 534
:::|||||

RESULT 9

A99708
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 01-Mar-2002
C;Accession: A99708
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99708
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA834056.1; PID:G13360091; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 050952
C;Genetics:
A;Gene: EC80633
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 55.0%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAAR 14
:::|||||
Db 521 VDKKQLRQWLASRA 534
:::|||||

RESULT 10

E85558
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
C;Accession: E85558
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <STO>
A;Cross-references: GB:AE005174; NID:G12513487; PIDN:AAG54929.1; GSPDB:GN00145; UWGP:Z0
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: entE
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 55.0%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAAR 14
:::|||||
Db 521 VDKKQLRQWLASRA 534
:::|||||

RESULT 10
E85558
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
C;Accession: E85558
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <STO>
A;Cross-references: GB:AE005174; NID:G12513487; PIDN:AAG54929.1; GSPDB:GN00145; UWGP:Z0
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: entE
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 55.0%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAAR 14
:::|||||
Db 521 VDKKQLRQWLASRA 534
:::|||||

RESULT 10
E85558
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
C;Accession: E85558
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <STO>
A;Cross-references: GB:AE005174; NID:G12513487; PIDN:AAG54929.1; GSPDB:GN00145; UWGP:Z0
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: entE
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 55.0%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAAR 14
:::|||||
Db 521 VDKKQLRQWLASRA 534
:::|||||

RESULT 10
E85558
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
C;Accession: E85558
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <STO>
A;Cross-references: GB:AE005174; NID:G12513487; PIDN:AAG54929.1; GSPDB:GN00145; UWGP:Z0
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: entE
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 55.0%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAAR 14
:::|||||
Db 521 VDKKQLRQWLASRA 534
:::|||||

RESULT 11
H83962
serine/threonine protein kinase BH2504 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: H83962
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11056132
A:Accession: H83962
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA06223.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2504

Query Match 55.0%; Score 44; DB 2; Length 664;
Best Local Similarity 46.7%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAARAP 15
:|||||: :
DB 90 VEGPTLKELIQGRP 104
:|||||: :
RESULT 12
H83415
cys/trans isomerase PA1846 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83415
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83415
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-762 <STO>
A:Cross-references: GB:AE004610; GB:AE004091; NID:g9947825; PIDN:AAG05235.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: Cti; PA1846

Query Match 55.0%; Score 44; DB 2; Length 762;
Best Local Similarity 64.3%; Pred. No. 28;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 EGPTLRQWLAAARAP 15
:|||||: :
DB 173 EYATLQKWLAAARAP 186
:|||||: :
RESULT 13
T44257
thiamine-phosphate diphosphorylase (EC 2.5.1.3) [imported] - Rhizobium etli plasmid b
C:Species: Rhizobium etli
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 03-Jun-2002
C:Accession: T44257
R:Miranda-Rios, J.; Morera, C.; Taboada, H.; Davalos, A.; Encarnacion, S.; Mora, J.; Sob
J. Bacteriol. 179, 6887-6893, 1997
A>Title: Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiotic
A:Reference number: 222737; MUID:98037482; PMID:9371431
A:Accession: T44257
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-204 <MIR>
A:Cross-references: EMBL:AF004408; NID:g2627325; PIDN:AAC45975.1; PID:g2627329

A:Experimental source: strain CE3
C:Genetics:
A:Gene: thiE
A:Genome: plasmid b
C:Superfamily: thiE protein; thiamin-phosphate pyrophosphorylase homology
C:Keywords: transferase

Query Match 53.8%; Score 43; DB 2; Length 204;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 7 ROWLAARAP 15
:|||||: :
DB 191 ROWLAATAP 199
:|||||: :
RESULT 14
G71337
probable dimethyladenosine transferase (ksgA) - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: G71337
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwir
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: G71337
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-285 <COL>
A:Cross-references: GB:AE001213; GB:AE000520; NID:g3322606; PIDN:AAC65323.1; PID:g332261
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0337
C:Superfamily: rRNA (adenine-N6-) methyltransferase

Query Match 53.8%; Score 43; DB 2; Length 285;
Best Local Similarity 64.3%; Pred. No. 15;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAARA 14
:|||||: :
DB 98 IEGDVLQWHAARA 111
:|||||: :
RESULT 15
G89894
protein kinase [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89894
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G89894
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <XUR>
A:Cross-references: GB:BA000018; PID:g13701020; PIDN:BA042315.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1063

Query Match 53.8%; Score 43; DB 2; Length 664;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAAARAP 15

Db ||||| : : : |
 90 IEGPTLSEYIESHGP 104

Search completed: June 24, 2003, 12:49:53
Job time : 17.1538 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:36:58 ; Search time 7.84615 Seconds
(without alignments)
79.293 Million cell updates/sec

Title: US-10-006-593-2

Perfect score: 80
Sequence: 1 IEGPTLRQWLARAP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

hed: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.5	61.9	333	1 CBBR_XANFL	P25545 xanthobacte
2	46	57.5	1242	1 NPHN_MOUSE	Q9qz87 mus musculu
3	45	56.2	1234	1 NPHN_RAT	Q9r044 rattus norv
4	44	55.0	536	1 ENTE_ECOLI	Q8xbv3 escherichia
5	44	55.0	536	1 ENTE_ECOLI	P10378 escherichia
6	43	53.8	204	1 THIE_RHET	Q34294 rhizobium e
7	43	53.8	825	1 SE5_EAT	P63003 rattus norv
8	42	52.5	1019	1 POL_SIV54	P12502 simian immu
9	41	51.2	195	1 ACPT_ECO57	Q8x504 escherichia
10	41	51.2	195	1 ACPT_ECOLI	P37623 escherichia
11	41	51.2	326	1 G3PC_TOBAC	P09094 nicotiana t
12	41	51.2	336	1 G3PC_PETCR	P26519 petroselinu
13	41	51.2	337	1 G3PC_MESCR	P17878 mesembryant
14	41	51.2	337	1 G3PC_SINAL	P04796 sinapis alb
15	41	51.2	338	1 G3PC_ARATH	P25858 arabidopsis
16	41	51.2	338	1 G3PC_DIACA	P34921 dianthus ca
17	41	51.2	399	1 PKNL_MYCTU	O53510 mycobacteri
18	40	50.0	203	1 INFB_CHICK	Q08773 gallus galli
19	40	50.0	463	1 YKAA_BACFI	P30267 bacillus fi
20	40	50.0	1028	1 EP42_HUMAN	P16452 homo sapien
21	40	50.0	1028	1 BGAL_ENTCL	Q47077 enterobacte
22	40	50.0	1712	1 CA24_HUMAN	P08572 homo sapien
23	39	48.8	267	1 BAF_BORPE	Q45338 bordetella
24	39	48.8	330	1 G3PC_LEIME	Q01558 leishmania
25	39	48.8	337	1 G3PC_PHACH	Q01982 phanerochaete
26	39	48.8	1022	1 POL_SIV5P	P19505 simian immu
27	39	48.8	1034	1 POL_HV2CA	P24107 human immu
28	39	48.8	1036	1 POL_HV2RO	P04584 human immu
29	39	48.8	1049	1 POL_HV2G1	P18042 human immu
30	39	48.8	1055	1 POL_HV2ST	P20876 human immu
31	39	48.8	1058	1 POL_HV2D2	P15833 human immu
32	38	47.5	297	1 XERC_MYCLE	Q9cbu0 mycobacteri
33	38	47.5	311	1 LYSR_ECOLI	P03030 escherichia

RESULT 1
CBBR_XANFL STANDARD; PRT; 333 AA.
ID CBBR_XANFL
AC P25545;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RUBISCO operon transcriptional regulator.
DE CBBR OR CFXO.
OS Xanthobacter flavus.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hyphomicrobium group; Xanthobacter.
OX NCBI_TaxID=281;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H4-14;
RX MEDLINE=94012468; PubMed=8407781;
RA van den Bergh E., Dijkhuizen L., Meijer W.G.;
RT "CbbR, a LysR-type transcriptional activator, is required for
RT expression of the autotrophic CO2 fixation enzymes of Xanthobacter
RT flavus";
RL J. Bacteriol. 175:6097-6104(1993).
RN [2]
RP SEQUENCE OF 1-150 FROM N.A.
RC STRAIN=H4-14;
RX MEDLINE=91172133; PubMed=1900916;
RA Meijer W.G., Arnberg A.C., Enequist H.G., Terpstra P., Lidstrom M.E.,
RA Dijkhuizen L.;
RT "Identification and organization of carbon dioxide fixation genes in
RT Xanthobacter flavus H4-14";
RL Mol. Gen. Genet. 225:320-330(1991).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR FOR THE CBB OPERON (CBBLSXFP)
CC FOR RUBISCO AND OTHER CALVIN CYCLE GENES. BINDS SPECIFICALLY TO
CC TWO BINDING SITES IN THE CBBR-CBBL INTERGENIC REGION.
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Z22705; CAA80406.1;
DR EMBL; X17252; -; NOT_ANNOTATED_CDS.
DR PIR; S13578; S13578.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
DR Transcription regulation; Activator; 1.
FT DNA BIND 22 41 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 333 AA; 36003 MW; 9B375B4FB2D1EE73 CRC64;

Q50652 mycobacteri
P25861 antirrhinum
P08735 zea mays (m
P26530 petunia hyb
Q09054 zea mays (m
P09317 ustilago ma
P17802 escherichia
Q8zyf6 pyrobaculum
P34783 atropis nu
P40115 sinapis alb
Q96524 arabidopsis
Q07700 mycobacteri

ALIGNMENTS

```

Query Match          61.9%; Score 49.5; DB 1; Length 333;
Best Local Similarity 66.7%; Pred. No. 0.4;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY      1 IEG-PTLROWLAARA 14
       :|||:|||||
Db      264 VEGLPVVRQWLAVRA 278

RESULT 2
NPNH MOUSE
ID NPNH MOUSE STANDARD; PRT; 1242 AA.
AC Q90ZS7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nephtrin precursor (Renal glomerulus-specific cell adhesion receptor).
GN NPHS1 OR NPNH
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.; AND SUBCELLULAR LOCATION.
RX MEDLINE=99436348; PubMed=10504495;
RA Holzman L.B.; St John P.L.; Kovari I.A.; Verma R.; Holthofer H.;
RA Abrahamson D.R.;
RT "Nephtrin localizes to the slit pore of the glomerular epithelial
cell.";
RL Kidney Int. 56:1481-1491(1999).
RN [2]
RP INTERACTION WITH CD2AP.
RX MEDLINE=21590051; PubMed=11733379;
RA Shih N.Y.; Li J.; Cotran R.; Mundel P.; Miner J.H.; Shaw A.S.;
RA "CD2AP localizes to the slit diaphragm and binds to nephrin via a
novel C-terminal domain.";
RT Am. J. Pathol. 159:2303-2308(2001).
RN [3]
RP INTERACTION WITH CD2AP AND NPHS2.
RX MEDLINE=21590460; PubMed=11733557;
RA Schwarz K.; Simons M.; Reiser J.; Saleem M.A.; Paul C.; Kriz W.;
RA Shaw A.S.; Holzman L.B.; Mundel P.;
RT "Podocin, a raft-associated component of the glomerular slit
diaphragm, interacts with CD2AP and nephrin.";
RL J. Clin. Invest. 108:1621-1629(2001).
CC -!- FUNCTION: Seems to play a role in the development or function of
the kidney glomerular filtration barrier. May anchor the podocyte
slit diaphragm to the actin cytoskeleton.
CC -!- SUBUNIT: Interacts with podocin/NPHS2 and with CD2AP C-terminal
domain.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential). Located
at podocyte slit diaphragm between podocyte foot processes.
CC -!- TISSUE SPECIFICITY: Expressed in kidney glomeruli.
CC -!- PTM: Phosphorylated on tyrosine residues (By similarity).
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 8 IMMUNOGLOBULIN-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF168466; AAF03368.1; -
CC MGD; MGI:1859637; Nphs1.
CC InterPro; IPR003961; FN.III.
CC InterPro; IPR003006; IG.MHC.
CC InterPro; IPR003598; IG.C2.
CC InterPro; IPR003600; IG_Like.

```

```

InterPro: IPR000399; TPP_enzyme.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; IG; 8.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_Like; 3.
KW Cell adhesion; Transmembrane; Signal; Glycoprotein;
KW Immunoglobulin domain; Repeat; Phosphorylation.
FT SIGNAL 1 22
FT CHAIN 23 1242
FT DOMAIN 23 1064 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1065 1086 POTENTIAL.
FT DOMAIN 1087 1242 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 46 118 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 153 224 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 258 324 IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 354 424 IG-LIKE C2-TYPE DOMAIN 4.
FT DOMAIN 458 535 IG-LIKE C2-TYPE DOMAIN 5.
FT DOMAIN 560 630 IG-LIKE C2-TYPE DOMAIN 6.
FT DOMAIN 754 823 IG-LIKE C2-TYPE DOMAIN 7.
FT DOMAIN 856 927 IG-LIKE C2-TYPE DOMAIN 8.
FT DOMAIN 941 1025 FIBRONECTIN TYPE-III.
FT DISULFID 53 111 POTENTIAL.
FT DISULFID 160 217 POTENTIAL.
FT DISULFID 285 317 POTENTIAL.
FT DISULFID 361 417 POTENTIAL.
FT DISULFID 465 528 POTENTIAL.
FT DISULFID 567 623 POTENTIAL.
FT DISULFID 761 816 POTENTIAL.
FT DISULFID 863 920 POTENTIAL.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 577 577 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 680 680 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 708 708 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 908 908 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1242 AA; 134890 MW; 02D82180BF145092 CRC64;

Query Match          57.5%; Score 46; DB 1; Length 1242;
Best Local Similarity 46.7%; Pred. No. 6;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 IEGPTLROWLAARAP 15
       :|||:|:|
Db      44 VEGSTIKLWCGVRAP 58

RESULT 3
NPNH RAT
ID NPNH RAT STANDARD; PRT; 1234 AA.
AC Q9R044; Q9QXX7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nephtrin precursor (Renal glomerulus-specific cell adhesion receptor).
GN NPHS1 OR NPNH
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX MEDLINE=99419288; PubMed=10487848;
RA Ahola H.; Wang S.-X.; Luimula P.; Solin M.-L.; Holzman L.B.;
RA Holthofer H.;
RT "Cloning and expression of the rat nephrin homolog.";
RL Am. J. Pathol. 155:907-913(1999).
RN [2]

```


SEQUENCE FROM N.A. (ISOFORM 1), AND SUBCELLULAR LOCATION.

RC STRAIN=Wistar; Tissue=Renal glomerulus;
 RX MEDLINE=20253275; PubMed=10792613;
 RA Kawachi H., Koike H., Kurihara H., Yaoita E., Orikasa M., Shia M.A.,
 RA Sakai T., Yamamoto T., Salant D.J., Shimizu F.;
 RT "Cloning of rat nephrin: expression in developing glomeruli and in
 RT proteinuric states.";
 RL Kidney Int. 57:1949-1961(2000).
 RN [3]

RP FUNCTION.
 RX MEDLINE=21868269; PubMed=11880318;
 RA Yuan H., Takeuchi E., Salant D.J.;
 RT "Podocyte slit-diaphragm protein nephrin is linked to the actin
 RT cytoskeleton.";
 RL Am. J. Physiol. 282:F585-F591(2002).
 CC -1- FUNCTION: Seems to play a role in the development or function of
 CC the kidney glomerular filtration barrier. May anchor the podocyte
 CC slit diaphragm to the cytoskeleton.
 CC -1- SUBUNIT: Interacts with podocin/NPHS2 and with CD2AP C-terminal
 CC domain (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential). Located
 CC at podocyte slit diaphragm between podocyte foot processes.
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms: isoform 1 (shown here), 2/alpha
 CC and 3/beta; may be produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Specifically expressed in the podocytes of
 CC kidney glomeruli.
 CC -1- PTM: Phosphorylated on tyrosine residues (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 8 IMMUNOGLOBULIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----

EMBL; AF125521; AAF12734.1; -;
 DR EMBL; AF161715; AAF14884.1; -;
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003598; IG_C2.
 DR InterPro; IPR003600; IG_like.
 DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00047; ig; 8.
 DR Pfam; PF00060; FN3; 1.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00410; IG_like; 6.
 DR Cell adhesion; Transmembrane; Signal; Glycoprotein;
 KW Immunglobulin domain; Repeat; Phosphorylation; Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 1234
 FT DOMAIN 19 1060
 FT TRANSMEM 1061 1081
 FT DOMAIN 1082 1234
 FT DOMAIN 42 114
 FT DOMAIN 149 220
 FT DOMAIN 254 320
 FT DOMAIN 350 420
 FT DOMAIN 454 531
 FT DOMAIN 556 626
 FT DOMAIN 750 819
 FT DOMAIN 852 923
 FT DOMAIN 938 1022
 FT DISULFID 49 107
 FT DISULFID 156 213
 FT DISULFID 261 313
 FT DISULFID 357 413
 FT DISULFID 461 524
 FT DISULFID 563 619
 FT DISULFID 757 812

FT DISULFID 859 916
 FT CARBOHYD 36 36
 FT CARBOHYD 352 352
 FT CARBOHYD 397 397
 FT CARBOHYD 543 543
 FT CARBOHYD 549 549
 FT CARBOHYD 560 560
 FT CARBOHYD 573 573
 FT CARBOHYD 704 704
 FT CARBOHYD 904 904
 FT VARSPLIC 1052 1064
 FT VARSPLIC 1052 1137
 FT CONFLICT 1 1
 FT CONFLICT 28 28
 FT CONFLICT 1211 1211
 FT SEQUENCE 1234 AA; 134304 MW; 6BF6707A229CA51E CRC64;
 Query Match 56.2%; Score 45; DB 1; Length 1234;
 Best Local Similarity 46.7%; Pred. No. 8.8;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 IEQPTLRQWLAAAP 15
 DB 40 VEGTTVKLMCGVRAP 54

RESULT 4
 ENTE ECO57 STANDARD; PRT; 536 AA.
 ID -ENTE_ECO57 STANDARD; PRT; 536 AA.
 AC Q9XBV3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enterobactin synthetase component E (Enterochelin synthase E)
 DE [includes: 2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58)
 DE (Dihydroxybenzoic acid-activating enzyme); S-
 DE dihydroxybenzoyltransferase (EC 2.3.1.-)].
 DE ENTE OR Z0736 OR ECS0633.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Ikeda T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: Activates the carboxylate group of 2,3-dihydroxy-
 CC benzoate (2,3-DHB), via ATP-dependent Ppi exchange reactions, to
 CC the acyladenylate. Then, catalyzes the acylation of holo-entB with
 CC 2,3-DHB adenylate, preparing that molecule for amide bond
 CC formation with L-serine (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + 2,3-dihydroxybenzoate = diphosphate +
 CC (2,3-dihydroxybenzoyl)-adenylate.

CC -!- CATALYTIC ACTIVITY: (2,3-dihydroxybenzoyl)-adenylate + holo-entB =
 CC adenosine 5'-monophosphate + acyl-holo-entB.
 CC -!- PATHWAY: Enterobactin biosynthesis.
 CC -!- SUBUNIT: Proteins entB, entD, entE, and entF form a multienzyme
 CC complex called enterobactin synthetase (by similarity).
 CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY. ENTE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE005239; AAC54929.1; -
 CC EMBL; AP002552; BAB34056.1; -
 CC InterPro: IPR000873; AMP-bind.
 CC Pfam: PF00501; AMP-binding; 1.
 CC PROSITE: PS00455; AMP BINDING; 1.
 CC Ligase; Transferase; Acyltransferase; Multifunctional enzyme;
 CC Enterobactin biosynthesis; Transport; Iron transport;
 CC Complete proteome.
 CC SEQUENCE 536 AA; 59040 MW; ABC80B3209940A5 CRC64;
 CC
 CC Query Match 55.0%; Score 44; DB 1; Length 536;
 CC Best Local Similarity 57.1%; Pred. No. 5.6;
 CC Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 IEQPTLRQWLAAARA 14
 CC : : |||||:
 CC 521 VDKKQLRWLASRA 534
 CC
 CC RESULT 5
 CC ENTE ECOLI STANDARD; PRT; 536 AA.
 CC ID ENTE ECOLI STANDARD; PRT; 536 AA.
 CC AC P10378; P15049; P77773;
 CC DT 01-MAR-1989 (Rel. 10, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Enterobactin synthetase component E (Enterochelin synthase E)
 CC DE [Includes: 2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58)
 CC DE (Dihydroxybenzoic acid-activating enzyme); S-
 CC DE dihydroxybenzoate transferase (EC 2.3.1.1).]
 CC GN ENTE OR B0594.
 CC OS Escherichia coli.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCBI_TaxID=562;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RC STRAIN=K12;
 CC RX MEDLINE=89290355; PubMed=2525505;
 CC RA Staab J.F., Elkins M.F., Earhart C.F.;
 CC RT "Nucleotide sequence of the Escherichia coli entE gene.";
 CC RL FEMS Microbiol. Lett. 50:15-19(1989).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC RC STRAIN=K12; MG1655;
 CC RX MEDLINE=97426617; PubMed=9278503;
 CC RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 CC RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 CC RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 CC RA Mau B., Shao Y.;
 CC RT "The complete genome sequence of Escherichia coli K-12.";
 CC RL Science 277:1232-1244(1997).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
 CC RA Federpspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
 CC RA Namath A., Oefner F., Roberts D., Schramm S., Davis R.W.;
 CC [4]
 CC Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC [4]
 CC SEQUENCE OF 1-8 FROM N.A.
 CC STRAIN=K12;
 CC RX MEDLINE=90236256; PubMed=2110093;
 CC RA Elkins M.F., Earhart C.F.;
 CC RT "Opacity factor from group A streptococci is an apoteinase.";
 CC RL FEMS Microbiol. Lett. 56:35-40(1988).
 CC [5]
 CC SEQUENCE OF 393-546 FROM N.A.
 CC RX MEDLINE=89123155; PubMed=2521622;
 CC Liu J., Duncan K., Walsh C.T.;
 CC RT "Nucleotide sequence of a cluster of Escherichia coli enterobactin
 CC biosynthesis genes: identification of entA and purification of its
 CC product 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase.";
 CC RL J. Bacteriol. 171:791-798(1989).
 CC [6]
 CC FUNCTION.
 CC RN PubMed=9214294;
 CC RX Gehring A.M., Bradley K.A., Walsh C.T.;
 CC RT "Enterobactin biosynthesis in Escherichia coli: isochorismate lyase
 CC (EntB) is a bifunctional enzyme that is phosphopantetheinylated by
 CC EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate.";
 CC RL Biochemistry 36:8495-8503(1997).
 CC -!- FUNCTION: Activates the carboxylate group of 2,3-dihydroxy-
 CC benzoate (2,3-DHB), via ATP-dependent P₁ exchange reactions, to
 CC the acyladenylate. Then, catalyzes the acylation of holo-entB with
 CC 2,3-DHB adenylate, preparing that molecule for amide bond
 CC formation with L-serine.
 CC -!- CATALYTIC ACTIVITY: ATP + 2,3-dihydroxybenzoate = diphosphate +
 CC (2,3-dihydroxybenzoyl)-adenylate.
 CC -!- CATALYTIC ACTIVITY: (2,3-dihydroxybenzoyl)-adenylate + holo-entB =
 CC adenosine 5'-monophosphate + acyl-holo-entB.
 CC -!- PATHWAY: Enterobactin biosynthesis.
 CC -!- SUBUNIT: Proteins entB, entD, entE, and entF form a multienzyme
 CC complex called enterobactin synthetase.
 CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY. ENTE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000165; AAC73695.1; -
 CC EMBL; U82598; AAB40794.1; -
 CC EMBL; X15058; CAA33158.1; -
 CC EMBL; M24148; AAA16101.1; -
 CC EMBL; M36700; AAA18492.1; -
 CC PIR; S08076; SYECES.
 CC PIR; J00078; J00078.
 CC PIR; A48308; A48308.
 CC EcoGene; EG10263; entE.
 CC InterPro: IPR000873; AMP-bind.
 CC Pfam: PF00501; AMP-binding; 1.
 CC PROSITE: PS00455; AMP BINDING; 1.
 CC Ligase; Transferase; Acyltransferase; Multifunctional enzyme;
 CC Enterobactin biosynthesis; Transport; Iron transport;
 CC Complete proteome.
 CC CONFLICT 369 378 DAEGNPLPQG -> ECRKRKSTAAR (IN REF. 1).
 CC SEQUENCE 536 AA; 59112 MW; F818942DFD8DC99 CRC64;
 CC
 CC Query Match 55.0%; Score 44; DB 1; Length 536;
 CC Best Local Similarity 57.1%; Pred. No. 5.6;
 CC Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 IEQPTLRQWLAAARA 14
 CC : : |||||:
 CC 521 VDKKQLRWLASRA 534
 CC
 CC Db

RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
protein 5ES in the nervous system.";
RL J. Biochem. 118:122-128(1995).
CC -!- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D37934; BAA07153.1; -
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86831 MW; AF667FE2FD555BDF CRC64;

Query Match 53.8%; Score 43; DB 1; Length 825;
Best Local Similarity 56.2%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 2 EGPT--LROWLAARAP 15
DB 88 EGPSLPLRQWLLPQCP 103
|||: ||||: |
POL_SIVS4
ID POL_SIVS4 STANDARD; PRT; 1019 AA.
AC P12502;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POL polyprotein (Contains: Protease (Retropepsin) (EC 3.4.23.-);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)).
GN POL.
OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11737;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89262053; PubMed=2786147;
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,
RA Johnson P.R.;
RT "An African primate lentivirus (SIVsm) closely related to HIV-2.";
RL Nature 339:389-392(1989).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonocester.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE
CC DETERMINED.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14307; -; NOT_ANNOTATED_CDS.
DR HSSP; P04584; 1PHV.
DR HIV; X14307; POL\$SMH4.
DR MEROPS; A02.002; -.
RN [1]

RC STRAIN=CE3;
RX MEDLINE=98037482; PubMed=9371431;
RA Miranda-Rios J., Morera C., Taboada H., Davalos A., Encarnacion S.,
RA Mora J., Soberon M.;
RT "Expression of thiamin biosynthetic genes (thiCOGE) and production of
RT symbiotic terminal oxidase cbb3 in Rhizobium etli.";
RL J. Bacteriol. 179:6887-6893(1997).
CC -!- FUNCTION: CONDENSES 4-METHYL-5-(BETA-HYDROXYETHYL)-THIAZOLE
CC MONOPHOSPHATE (THZ-P) AND 4-AMINO-5-HYDROXYMETHYL PRIMIDINE
CC PYROPHOSPHATE (HMP-PP) TO FORM THIAMINE MONOPHOSPHATE (TMP) (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2-methyl-4-amino-5-hydroxymethylpyrimidine
CC diphosphate + 4-4-methyl-5-(2-phosphonooxyethyl)-thiazole =
CC diphosphate + thiamine monophosphate.
CC -!- PATHWAY: Thiamine biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE TMP-PPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF004408; AAC45975.1; -
DR HSSP; P39594; 2TPS.
DR InterPro; IPR003733; TMP synthase.
PRam; PF02581; TMP-TEN1; 1.
Thiamine biosynthesis; Transferase; Plasmid.
DOMAIN 46 54 POLY-ALA.
SQ SEQUENCE 204 AA; 22488 MW; 81BA95165880628A CRC64;

Query Match 53.8%; Score 43; DB 1; Length 204;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ROWLAARAP 15
DB 191 ROWLAATAP 199
|||: ||||: ||
POL_SIVS4
ID 5P5 RAT STANDARD; PRT; 825 AA.
AC O63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 5E5 antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RC STRAIN=CE3;
RX MEDLINE=98037482; PubMed=9371431;
RA Miranda-Rios J., Morera C., Taboada H., Davalos A., Encarnacion S.,
RA Mora J., Soberon M.;
RT "Expression of thiamin biosynthetic genes (thiCOGE) and production of
RT symbiotic terminal oxidase cbb3 in Rhizobium etli.";
RL J. Bacteriol. 179:6887-6893(1997).
CC -!- FUNCTION: CONDENSES 4-METHYL-5-(BETA-HYDROXYETHYL)-THIAZOLE
CC MONOPHOSPHATE (THZ-P) AND 4-AMINO-5-HYDROXYMETHYL PRIMIDINE
CC PYROPHOSPHATE (HMP-PP) TO FORM THIAMINE MONOPHOSPHATE (TMP) (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2-methyl-4-amino-5-hydroxymethylpyrimidine
CC diphosphate + 4-4-methyl-5-(2-phosphonooxyethyl)-thiazole =
CC diphosphate + thiamine monophosphate.
CC -!- PATHWAY: Thiamine biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE TMP-PPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF004408; AAC45975.1; -
DR HSSP; P39594; 2TPS.
DR InterPro; IPR003733; TMP synthase.
PRam; PF02581; TMP-TEN1; 1.
Thiamine biosynthesis; Transferase; Plasmid.
DOMAIN 46 54 POLY-ALA.
SQ SEQUENCE 204 AA; 22488 MW; 81BA95165880628A CRC64;

Query Match 53.8%; Score 43; DB 1; Length 204;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ROWLAARAP 15
DB 191 ROWLAATAP 199
|||: ||||: ||
POL_SIVS4
ID 5P5 RAT STANDARD; PRT; 825 AA.
AC O63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 5E5 antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RC STRAIN=CE3;
RX MEDLINE=98037482; PubMed=9371431;
RA Miranda-Rios J., Morera C., Taboada H., Davalos A., Encarnacion S.,
RA Mora J., Soberon M.;
RT "Expression of thiamin biosynthetic genes (thiCOGE) and production of
RT symbiotic terminal oxidase cbb3 in Rhizobium etli.";
RL J. Bacteriol. 179:6887-6893(1997).
CC -!- FUNCTION: CONDENSES 4-METHYL-5-(BETA-HYDROXYETHYL)-THIAZOLE
CC MONOPHOSPHATE (THZ-P) AND 4-AMINO-5-HYDROXYMETHYL PRIMIDINE
CC PYROPHOSPHATE (HMP-PP) TO FORM THIAMINE MONOPHOSPHATE (TMP) (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2-methyl-4-amino-5-hydroxymethylpyrimidine
CC diphosphate + 4-4-methyl-5-(2-phosphonooxyethyl)-thiazole =
CC diphosphate + thiamine monophosphate.
CC -!- PATHWAY: Thiamine biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE TMP-PPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF004408; AAC45975.1; -
DR HSSP; P39594; 2TPS.
DR InterPro; IPR003733; TMP synthase.
PRam; PF02581; TMP-TEN1; 1.
Thiamine biosynthesis; Transferase; Plasmid.
DOMAIN 46 54 POLY-ALA.
SQ SEQUENCE 204 AA; 22488 MW; 81BA95165880628A CRC64;

Query Match 53.8%; Score 43; DB 1; Length 204;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ROWLAARAP 15
DB 191 ROWLAATAP 199
|||: ||||: ||
POL_SIVS4
ID 5P5

RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
protein 5ES in the nervous system.";
RL J. Biochem. 118:122-128(1995).
CC -!- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D37934; BAA07153.1; -
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86831 MW; AF667FE2FD555BDF CRC64;

Query Match 53.8%; Score 43; DB 1; Length 825;
Best Local Similarity 56.2%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 2 EGPT--LROWLAARAP 15
DB 88 EGPSLPLRQWLLPQCP 103
|||: ||||: |

RESULT 8
POL_SIVS4
ID POL_SIVS4 STANDARD; PRT; 1019 AA.
AC P12502;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POL polyprotein (Contains: Protease (Retropepsin) (EC 3.4.23.-);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)).
GN POL.
OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11737;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89262053; PubMed=2786147;
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,
RA Johnson P.R.;
RT "An African primate lentivirus (SIVsm) closely related to HIV-2.";
RL Nature 339:389-392(1989).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE
CC DETERMINED.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14307; -; NOT_ANNOTATED_CDS.
DR HSSP; P04584; 1PHV.
DR HIV; X14307; POL\$SMH4.
DR MEROPS; A02.002; -.
DR [1]

RC STRAIN=CE3;
RX MEDLINE=98037482; PubMed=9371431;
RA Miranda-Rios J., Morera C., Taboada H., Davalos A., Encarnacion S.,
RA Mora J., Soberon M.;
RT "Expression of thiamin biosynthetic genes (thiCOGE) and production of
RT symbiotic terminal oxidase cbb3 in Rhizobium etli.";
RL J. Bacteriol. 179:6887-6893(1997).
CC -!- FUNCTION: CONDENSES 4-METHYL-5-(BETA-HYDROXYETHYL)-THIAZOLE
CC MONOPHOSPHATE (THZ-P) AND 4-AMINO-5-HYDROXYMETHYL PRIMIDINE
CC PYROPHOSPHATE (HMP-PP) TO FORM THIAMINE MONOPHOSPHATE (TMP) (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2-methyl-4-amino-5-hydroxymethylpyrimidine
CC diphosphate + 4-4-methyl-5-(2-phosphonoxyethyl)-thiazole =
CC diphosphate + thiamine monophosphate.
CC -!- PATHWAY: Thiamine biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE TMP-PPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF004408; AAC45975.1; -
DR HSSP; P39594; 2TPS.
DR InterPro; IPR003733; TMP synthase.
PRam; PF02581; TMP-TEN1; 1.
Thiamine biosynthesis; Transferase; Plasmid.
DOMAIN 46 54 POLY-ALA.
SQ SEQUENCE 204 AA; 22488 MW; 81BA95165880628A CRC64;

Query Match 53.8%; Score 43; DB 1; Length 204;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ROWLAARAP 15
DB 191 ROWLAATAP 199
|||: ||||: ||

RESULT 7
5ES_RAT
ID 5ES_RAT STANDARD; PRT; 825 AA.
AC O63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 5ES antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
protein 5ES in the nervous system.";
RL J. Biochem. 118:122-128(1995).
CC -!- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D37934; BAA07153.1; -
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86831 MW; AF667FE2FD555BDF CRC64;

Query Match 53.8%; Score 43; DB 1; Length 825;
Best Local Similarity 56.2%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 2 EGPT--LROWLAARAP 15
DB 88 EGPSLPLRQWLLPQCP 103
|||: ||||: |

RESULT 8
POL_SIVS4
ID POL_SIVS4 STANDARD; PRT; 1019 AA.
AC P12502;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POL polyprotein (Contains: Protease (Retropepsin) (EC 3.4.23.-);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)).
GN POL.
OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11737;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89262053; PubMed=2786147;
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,
RA Johnson P.R.;
RT "An African primate lentivirus (SIVsm) closely related to HIV-2.";
RL Nature 339:389-392(1989).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE
CC DETERMINED.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14307; -; NOT_ANNOTATED_CDS.
DR HSSP; P04584; 1PHV.
DR HIV; X14307; POL\$SMH4.
DR MEROPS; A02.002; -.
DR [1]

DR InterPro: IPR001995; Aspartate rtrv.
 DR InterPro: IPR001969; Aspartate rtrv.
 DR InterPro: IPR001037; Integrase C.
 DR InterPro: IPR003308; Integrase Zn.
 DR InterPro: IPR002156; RNaseH.
 DR InterPro: IPR000477; RVase.
 DR InterPro: IPR001584; Rve.
 DR Pfam: PF000075; rnaaseh.1.
 DR Pfam: PF000078; rvp.1.
 DR Pfam: PF000078; rvt.1.
 DR Pfam: PF000532; integrase.1.
 DR Pfam: PF00665; rve.1.
 DR Pfam: PF02022; Integrase Zn.1.
 DR PROSITE: PS00141; ASP PROTEASE.1.
 DR PROSITE: PS00175; ASP-PROT RETROV.1.
 KW AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
 KW Nuclease; Transferase; RNA-directed DNA polymerase.
 FT CHAIN 1 167
 FT AC_SITE 93 93
 SQ SEQUENCE 1019 AA; 115465 MW; 8D3DE0B85FC92B1C CRC64;

Query Match 52.5%; Score 42; DB 1; Length 1019;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
 DB 184 EGPKLRQW 191

RESULT 9
 ACPT ECO57
 ID ACPT ECO57 STANDARD; PRT; 195 AA.
 AC Q8X5U4;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 4'-phosphopantetheinyl transferase acpt (EC 2.7.8.-).
 GN ACPT OR 24867 OR ECS4342.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OC NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11205551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoculis K.,
 Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 Welch R.A., Blattner F.R.;
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22 (2001).
 CC -!- FUNCTION: May be involved in an alternative pathway for
 phosphopantetheinyl transfer and holo-ACP synthesis in E.coli. The
 native apo-protein substrate is unknown. Is able to functionally
 replace acpS in vivo but only when expressed at high levels (By
 similarity)
 CC -!- CATALYTIC ACTIVITY: CoA + [protein-X] = adenosine 3',5'-
 bisphosphate + phosphopantetheinyl-[protein-X]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 GSP/SFP/HETI/ACPT FAMILY.

CC -!- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY.
 CC GSP/SFP/HETI/ACPT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AE005571; AAG58602.1; -
 DR EMBL: AE005571; AAG58602.1; -
 DR InterPro: IPR002582; ACPs.
 DR Pfam: PF01648; ACPs; 1
 KW Transferase; Complete proteome.
 SQ SEQUENCE 195 AA; 21737 MW; 360557D7230B2AB7 CRC64;

Query Match 51.2%; Score 41; DB 1; Length 195;
 Best Local Similarity 53.8%; Pred. No. 6.5;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAAPA 14
 DB 27 QGPRRERWLAGRA 39

RESULT 10
 ACPT ECOLI
 ID ACPT ECOLI STANDARD; PRT; 195 AA.
 AC P37623;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 4'-phosphopantetheinyl transferase acpt (EC 2.7.8.-).
 GN ACPT OR B3475.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MGI1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Buxland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RA "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes";
 RL Nucleic Acids Res. 22:2576-2586 (1994).
 RN [2]
 RP FUNCTION.
 RC PubMed=8939709;
 RA Lambalot R.H., Gehring A.M., Flugel R.S., Zuber P., LaCelle M.,
 Marahiel M.A., Reid R., Khosla C., Walsh C.T.;
 RA "A new enzyme superfamily - the phosphopantetheinyl transferases";
 RL Chem. Biol. 3:923-936 (1996).
 RN [3]
 RP FUNCTION.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=20092861; PubMed=10625633;
 RA Flugel R.S., Hwangbo Y., Lambalot R.H., Cronan J.E. Jr., Walsh C.T.;
 RA "Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer
 in Escherichia coli";
 RL J. Biol. Chem. 275:959-968 (2000).
 CC -!- FUNCTION: May be involved in an alternative pathway for
 phosphopantetheinyl transfer and holo-ACP synthesis in E.coli. The
 native apo-protein substrate is unknown. Is able to functionally
 replace acpS in vivo but only when expressed at high levels.
 CC -!- CATALYTIC ACTIVITY: CoA + [protein-X] = adenosine 3',5'-
 bisphosphate + phosphopantetheinyl-[protein-X]
 CC -!- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY.
 CC GSP/SFP/HETI/ACPT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U00039; AAB18450.1; -
DR EMBL; AE000423; AAC76500.1; -
DR EcoGene; EG12221; acpt.
DR InterPro; IPR002582; ACPS.
DR Pfam; PF01648; ACPS; 1.
DR Transferase; Complete proteome.
KW SEQUENCE 195 AA; 21768 MW; 29385FDA343B2AB7 CRC64;
SQ

Query Match 51.2%; Score 41; DB 1; Length 195;
Best Local Similarity 53.8%; Pred. No. 6.5;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
DB 2 EGPTLRWLAAARA 14
:|||||
27 QGPRRERWLAGRA 39

RESULT 11
G3PC TOBAC STANDARD; PRT; 326 AA.
AC P09094;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE (Fragment).
GN GAPC.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87002494; PubMed=3757034;
RT Shih M.-C.; Lazar G.; Goodman H.M.;
RT "Evidence in favor of the symbiotic origin of chloroplasts: primary
RT structure and evolution of tobacco glyceraldehyde-3-phosphate
RT dehydrogenases".
RT Cell 47:73-80(1986).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
CC WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE
CC ENCODED BY DISTINCT GENES.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14419; AAA34077.1; -
DR PIR; C24430; C24430.
DR HSSP; P00357; 4GPD.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PF02800; gpdh; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family.
FT BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 336 AA; 36371 MW; 9C052A143AFA2494 CRC64;
Query Match 51.2%; Score 41; DB 1; Length 336;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEGPTLRWLAAARA 14
:|||||

KW Glycolysis; Oxidoreductase; NAD; Multigene family.
FT NON_TER 1
FT BINDING 143 143 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 170 170 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 326 AA; 35533 MW; A3PFD7D1AA91038A CRC64;
Query Match 51.2%; Score 41; DB 1; Length 326;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEGPTLRWLAAARA 14
:|||||
179 VDGPMSKDWGRGA 192
DB

RESULT 12
G3PC PETCR STANDARD; PRT; 336 AA.
AC P26519;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE GAPC OR GAPDH.
GN Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
OX NCBI_TaxID=4043;
RN [1]
RP SEQUENCE FROM N.A.
RX Martin W.; Giehl A.; Saedler H.;
RT "Molecular evidence for pre-Cretaceous angiosperm origins".
RT Nature 339:46-48(1989).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC
CC FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
CC WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE
CC ENCODED BY DISTINCT GENES.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60344; CAA42902.1; -
DR PIR; S18484; DEPZG.
DR HSSP; P00357; 4GPD.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PF02800; gpdh; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family.
FT BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 336 AA; 36371 MW; 9C052A143AFA2494 CRC64;
Query Match 51.2%; Score 41; DB 1; Length 336;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEGPTLRWLAAARA 14
:|||||

```

Db      189 VDGFSMKDWRGGA 202

RESULT 13
G3PC_MESCR
ID_G3PC_MESCR STANDARD; PRT; 337 AA.
AC      P17878;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12).
GN      G3PC.
OS      Sinapis alba (White mustard) (Brassica hirta).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Sinapis.
OX      NCBI_TaxID=3728;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90154012; PubMed=2303458;
RA      Ostrem J.A., Vernon D.M., Bohnert H.J.;
RT      "Increased expression of a gene coding for NAD:glyceraldehyde-3-
RT      phosphate dehydrogenase during the transition from C3 photosynthesis
RT      to crassulacean acid metabolism in Mesembryanthemum crystallinum.";
RL      J. Biol. Chem. 265:3497-3502(1990).
CC      -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC      NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC      -!- PATHWAY: Second phase of glycolysis; first step.
CC      -!- SUBUNIT: HOMOTETRAMER.
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -!- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC
CC      FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
CC      WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE
CC      ENCODED BY DISTINCT GENES.
CC      -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC      DEHYDROGENASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; J05223; AAA33033.1; -.
DR      EMBL; M29556; AAA33031.1; -.
DR      PIR; A35080; A35080.
DR      HSP; P00357; 4GPD.
DR      InterPro; IPR000173; GAP_dhhydrogenase.
DR      Pfam; PF00044; spdh_C; 1.
DR      PRINTS; PR00078; G3PDHGRGNASE.
DR      PROSITE; PS00071; GAPDH; 1.
KW      Glycolysis; Oxidoreductase; NAD; Multigene family.
FT      DOMAIN 1 152
FT      DOMAIN 153 337
FT      DOMAIN 155 74
FT      DOMAIN 182 205
FT      SIMILAR 101 337
FT      BINDING 155 155
FT      ACT SITE 182 182
FT      BINDING 155 155
FT      ACT SITE 182 182
SQ      SEQUENCE 337 AA; 36587 MW; B07B28A2509E9EA CRC64;

Query Match 51.2%; Score 41; DB 1; Length 337;
Best Local Similarity 35.7%; Pred. NO. 11;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY      1 IEQPTLRQWLAARA 14
Db      190 VDGFSMKDWRGGA 203

RESULT 14
G3PC_SINAL
ID_G3PC_SINAL STANDARD; PRT; 337 AA.
AC      P17878;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12).
GN      G3PC.
OS      Sinapis alba (White mustard) (Brassica hirta).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Sinapis.
OX      NCBI_TaxID=3728;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90154012; PubMed=2303458;
RA      Ostrem J.A., Vernon D.M., Bohnert H.J.;
RT      "Increased expression of a gene coding for NAD:glyceraldehyde-3-
RT      phosphate dehydrogenase during the transition from C3 photosynthesis
RT      to crassulacean acid metabolism in Mesembryanthemum crystallinum.";
RL      J. Biol. Chem. 265:3497-3502(1990).
CC      -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC      NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC      -!- PATHWAY: Second phase of glycolysis; first step.
CC      -!- SUBUNIT: HOMOTETRAMER.
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -!- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC
CC      FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
CC      WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE
CC      ENCODED BY DISTINCT GENES.
CC      -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC      DEHYDROGENASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; J05223; AAA33033.1; -.
DR      EMBL; M29556; AAA33031.1; -.
DR      PIR; A35080; A35080.
DR      HSP; P00357; 4GPD.
DR      InterPro; IPR000173; GAP_dhhydrogenase.
DR      Pfam; PF00044; spdh_C; 1.
DR      PRINTS; PR00078; G3PDHGRGNASE.
DR      PROSITE; PS00071; GAPDH; 1.
KW      Glycolysis; Oxidoreductase; NAD; Multigene family.
FT      DOMAIN 1 151
FT      DOMAIN 152 337
FT      DOMAIN 154 154
FT      BINDING 154 154
FT      ACT SITE 181 181
FT      BINDING 154 154
FT      ACT SITE 181 181
SQ      SEQUENCE 337 AA; 36587 MW; B07B28A2509E9EA CRC64;

Query Match 51.2%; Score 41; DB 1; Length 337;
Best Local Similarity 35.7%; Pred. NO. 11;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

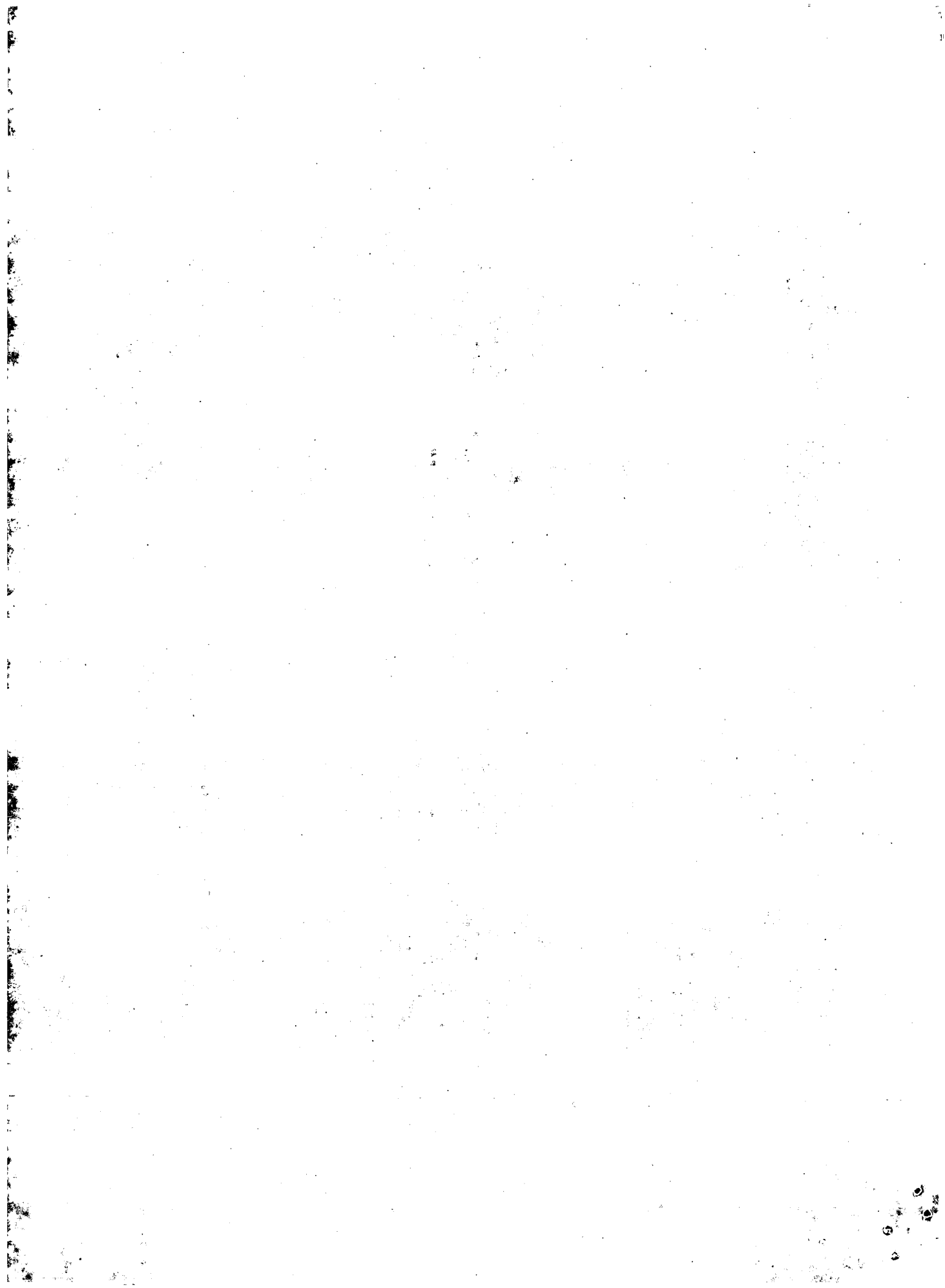
OY      1 IEQPTLRQWLAARA 14
Db      190 VDGFSMKDWRGGA 203

RESULT 15
G3PC_ARATH
ID_G3PC_ARATH STANDARD; PRT; 338 AA.
AC      P25858;
DT      01-MAY-1992 (Rel. 22, Created)

```

DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN GAPC OR AT3G04120 OR T6K12.26.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 ON NCBI_TaxID=3702;
 RP MEDLINE=93013005; PubMed=1916285;
 RX Shih M.-C., Heinrich P., Goodman H.M.;
 RA "Cloning and cytosomal mapping of nuclear genes encoding
 RT chloroplast and cytosolic glyceraldehyde-3-phosphate-dehydrogenase
 RT from Arabidopsis thaliana";
 RL Gene 104:133-138 (1991).
 RL [2]
 RP ERRATUM.
 RX MEDLINE=93013005; PubMed=1998114;
 RA Shih M.-C., Heinrich P., Goodman H.M.;
 RL Gene 119:317-319 (1992).
 RL [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salancubut M., Lemcke K., Rieger M., Ansoorge W., Unseid M.,
 RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
 RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
 RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,
 RA Wincker P., Catolico L., Weissbach J., Saurin W., Quetier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurmbach E., Drzonek H., Erff H., Jordan N., Bangert S.,
 RA Wiedelmann R., Kranz H., Hoffmann R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
 RA Conrad A., Horrischer K., Kauer G., Loehner T.-H., Nordsiek G.,
 RA Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
 RA Cooke R., Laudie M., Berger-Llauró C., Purnelle B., Masuy D.,
 RA de Haan M., Maarsee A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Matti R., Wu D., Peterson J., Van Aken S.,
 RA Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato I., Asamizu E.,
 RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 RL thaliana";
 RL Nature 408:820-822 (2000).
 RL [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RA "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGECC).";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RL [5]
 RP SEQUENCE OF 181-321 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX Cooke R., Laudie M., Raynal M., Delseny M.;
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RL CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -|- PATHWAY: Second phase of glycolysis; first step.
 CC -|- SUBUNIT: HOMOTETRAMER.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC

CC FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
 CC WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE
 CC ENCODED BY DISTINCT GENES.
 CC -|- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M641116; AAA32794.1; -;
 CC EMBL; M641119; AAA32796.1; -;
 CC EMBL; AC016829; AAF26801.1; -;
 CC EMBL; AY052267; AAK97737.1; -;
 CC EMBL; AY060521; AAL31134.1; -;
 CC EMBL; F20074; CAA23391.1; -;
 CC FIR; JQ1287; JQ1287.
 CC HSSP; P00357; 4GPD.
 CC InterPro: IPR000173; GAP_dhhdhrogenase.
 CC Pfam; PF00044; gpdh; 1.
 CC Pfam; PF02800; gpdh C; 1.
 CC PRINTS; PR00078; G3PDHGRNASE.
 CC PROSITE; PS00071; GAPDH; 1.
 CC BINDING 156 156 GLYCERALDEHYDE 3-PHOSPHATE.
 CC ACT_SITE 183 183 ACTIVATES THIOL GROUP DURING CATALYSIS.
 CC CONFLICT 127 127 A -> E (IN REF. 1).
 CC CONFLICT 260 260 D -> E (IN REF. 1).
 CC SEQUENCE 338 AA; 36914 MW; 4186F65E1F1EE96F CRC64;
 Query Match 51.2%; Score 41; DB 1; Length 338;
 Best Local Similarity 35.7%; Pred. No. 11;
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 IEQPTLRQWLAAARA 14
 DB 192 VDGPSPKMDWRGGA 205
 Search completed: June 24, 2003, 12:48:35
 Job time : 8.84615 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:44:08 ; Search time 33.4615 Seconds
(without alignments)
92.366 Million cell updates/sec

Title: US-10-006-593-2

Perfect score: 80
Sequence: 1 IEQPTLRQWLARAP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rviro:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	61.3	91	16	Q8Y015
2	49	61.3	319	16	Q9RKM5
3	48.5	60.6	607	2	Q9L8D4
4	48	60.0	252	16	Q8XPQ9
5	48	60.0	766	2	Q9WZ0
6	48	60.0	941	12	Q8QUJ6
7	47	58.8	296	16	Q8ZGS7
8	46	57.5	252	16	Q8XY68
9	46	57.5	306	16	Q05576
10	46	57.5	400	16	Q69568
11	46	57.5	766	2	Q9R9Y9
12	46	57.5	766	2	Q33466
13	45	57.5	1256	11	Q9JIX1
14	45	56.2	87	16	Q8XZR4
15	45	56.2	245	2	O66272
16	45	56.2	249	2	O82989

17	45	56.2	278	2	Q9XDV0
18	45	56.2	494	11	Q9ESC6
19	45	56.2	1049	2	Q9XBP6
20	45	56.2	1252	11	Q9JIX2
21	45	56.2	1256	11	Q9ET59
22	45	56.2	1256	11	Q9Z5S5
23	44.5	55.6	436	2	Q45293
24	44	55.0	207	5	O16116
25	44	55.0	536	16	Q8XBV3
26	44	55.0	664	16	Q9K9Z0
27	44	55.0	762	16	Q9I2P9
28	43.5	54.4	168	5	Q9V492
29	43	53.8	285	16	O83357
30	43	53.8	354	17	Q8ZYT5
31	43	53.8	388	2	Q9KX10
32	43	53.8	664	16	Q99UP8
33	43	53.8	683	16	O83436
34	43	53.8	754	5	Q9SY82
35	42	52.5	201	2	Q9Z3X4
36	42	52.5	313	15	P90433
37	42	52.5	326	2	P95613
38	42	52.5	352	16	Q9I763
39	42	52.5	367	16	O8Y030
40	42	52.5	438	16	Q9AA92
41	42	52.5	450	10	Q9SLB9
42	42	52.5	586	5	Q9NGP9
43	42	52.5	1019	15	P89154
44	42	52.5	2617	16	Q8YTR9
45	41	51.2	75	16	Q98AJ1

ALIGNMENTS

RESULT 1

Q8Y015 PRELIMINARY; PRT; 91 AA.
AC Q8Y015;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RSC1059.
GN RSC1059 OR RS04149.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646062; CAD14761.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 91 AA; 10321 MW; 2B4DFEB37A528AD CRC64;

Query Match 61.3%; Score 49; DB 16; Length 91;
Best Local Similarity 46.7%; Pred. No. 1.7;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEQPTLRQWLARAP 15

Db 75 LDGPAVQWLAAQTP 89

RESULT 2

Q9RKM5 ID Q9RKM5 PRELIMINARY; PRT; 319 AA.
AC Q9RKM5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative MerR family transcriptional regulator.
GN SCO4102 OR SC017.06C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1502;
RN 1 IEGPTLROWLAARAP 15
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
CC -1- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AL118515; CAB56383.1; --
DR InterPro; IPR000551; HTH_MerR.
DR Pfam; PF00376; merR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MERR; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;
Query Match 61.3%; Score 49; DB 16; Length 319;
Best Local Similarity 66.7%; Pred. No. 5.9;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 EGPTLROWLAAR 13
Db 258 DGPDLREWLGR 269
RESULT 3
Q9L8D4 ID Q9L8D4 PRELIMINARY; PRT; 607 AA.
AC Q9L8D4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Hypothetical 66.3 kDa protein (Fragment).

OS Polyangium cellulorum.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Sorangineae; Polyangiaceae; Polyangium.
OX NCBI_TaxID=56;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=SO CE90;
RX MEDLINE=20130945; PubMed=10662695;
RA Molnar I., Schupp T., Ono M., Zirkle R.E., Milnamow M., Cyr D.D.,
RA Nowak-Thompson B., Engel N., Toupet C., Stratmann A., Ligon J.M.,
RA Gorlach J., Mayo J.M., Hu A., Goff S., Schmid J., Ligon J.M.;
RT "The biosynthetic gene cluster for the microtubule-stabilizing agents
RT epothilones A and B from Sorangium cellulorum So ce90.";
RL Chem. Biol. 7:97-109(2000).
RL EMBL; AF210843; AAF26904.1; --
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 607 AA; 66326 MW; F113CA299B25048E CRC64;
Query Match 60.6%; Score 48.5; DB 2; Length 607;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
QY 1 IEGPTLROWLAAR-AP 15
Db 96 VDGPAVRLVLAARGAP 111
RESULT 4
Q8XPQ9 ID Q8XPQ9 PRELIMINARY; PRT; 252 AA.
AC Q8XPQ9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative transcription regulator protein.
GN RSP1579 OR RS02135.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Ralstonia solanacearum.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brotter P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigquier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:457-502(2002).
RL EMBL; AL646085; CAD18730.1; --
DR InterPro; IPR000792; HTH_LuxR.
DR Pfam; PF00196; GersE; 1.
DR PRINTS; PR00038; HTHLUXR.
DR SMART; SM00421; HTH_LUXR; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 252 AA; 27666 MW; 483403E326F7C2E CRC64;
Query Match 60.0%; Score 48; DB 16; Length 252;
Best Local Similarity 53.3%; Pred. No. 6.8;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEGPTLROWLAARAP 15
Db 76 IDTPIMRRLATRRP 90
RESULT 5
Q9WWZO

```

ID Q9WZ0 PRELIMINARY; PRT; 766 AA.
AC Q9WZ0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Putative membrane protein.
DE Cti.
DE Cti.
GN Cti.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DOT-TIE;
RX MEDLINE=99412268; PubMed=10482510;
RA Junker F., Ramos J.L.;
RT "Involvement of the cis/trans isomerase cti in solvent resistance of
RT Pseudomonas putida DOT-TIE."
RA J. Bacteriol. 181:5693-5700(1999).
RA EMBL; AF110738; AAB41252.1; -.
DR InterPro; IPR000345; CytC heme bind.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
SQ SEQUENCE 766 AA; 87058 MW; A4A0FC6C2C301FE CRC64;

Query Match 60.0%; Score 48; DB 2; Length 766;
Best Local Similarity 71.4%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAARAP 15
Db 177 EYATLRWLAAGAP 190

RESULT 6
Q8QUJ6 PRELIMINARY; PRT; 941 AA.
ID Q8QUJ6
AC Q8QUJ6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE ORF114L.
DE ORF114L.
OS infectious spleen and kidney necrosis virus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae;
OC Unclassified Iridoviridae.
OX NCBI_TaxID=180170;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21874810; PubMed=11878882;
RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,
RA Chan S.M.;
RT "Complete genome analysis of the mandarin fish infectious spleen and
RT kidney necrosis iridovirus."
RL Virology 291:126-139(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,
RA Chan S.M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF371960; AAL98838.1; -.
SQ SEQUENCE 941 AA; 106703 MW; EB663998CTF6CE83 CRC64;

Query Match 60.0%; Score 48; DB 12; Length 941;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
Db 581 VQPTLAQWICSTA 594

RESULT 7
Q8ZGS7 PRELIMINARY; PRT; 296 AA.
ID Q8ZGS7

```

```

AC Q8ZGS7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Putative membrane protein.
DE YPO1203.
GN YPO1203.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feitwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
DR EMBL; AJ414147; CAC90042.1; -.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 296 AA; 31378 MW; 45947413DCD54CFF CRC64;

Query Match 58.8%; Score 47; DB 16; Length 296;
Best Local Similarity 81.8%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PTLRQWLAARA 14
Db 66 PTLRQWLAARA 76

RESULT 8
Q8XY68 PRELIMINARY; PRT; 252 AA.
ID Q8XY68
AC Q8XY68;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Putative transcription regulator protein.
DE RSC1895 OR RS03457.
GN Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Denange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646067; CAD15597.1; -.
DR InterPro; IPR000792; HTH_LuxR.
DR Pfam; PF00196; GexE; 1.
DR PRINTS; PR00038; HTHLUXR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00421; HTH_LUXR; 1.
KW Complete proteome.
SQ SEQUENCE 252 AA; 27945 MW; 17FD89DC6803EBFF CRC64;

Query Match 57.5%; Score 46; DB 16; Length 252;

```

Best Local Similarity 46.7%; Pred. No. 14; Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLARAP 15
DB 76 IDTLPLMRWFATRSP 90

RESULT 9

OC 005576 PRELIMINARY; PRT; 306 AA.
AC 005576;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GALU (UTP-glucose-1-phosphate uridylyltransferase).
GN GALU OR RV0993 OR MT1022 OR MTC1237.07.
OS Mycobacterium tuberculosis
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=11773;
[1]
SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmier K., Gas S., Barry C.E. III, Tekai F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares R., Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
RT Nature 393:537-544 (1998).
RL [2]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z94752; CAB08153.1; -;
DR EMBL; AE006986; AAK45269.1; -;
TI TIGR; MT1022; -;
DR TubercuList; RV0993; -;
DR InterPro; IPR001825; NTP transferase.
DR Pfam; PF00483; NTP_transferase; 1.
KW Complete proteome.
FT CONFLICT 235 Q -> R (IN REF. 2).
SQ SEQUENCE 306 AA; 32378 MW; 24C2397443B0A3E8 CRC64;

Query Match 57.5%; Score 45; DB 16; Length 306;
Best Local Similarity 72.7%; Pred. No. 17; Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAR 13

DB 290 GFDLRRLVAR 300

RESULT 10

OC 069568 PRELIMINARY; PRT; 400 AA.
AC 069568;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Serine-threonine protein kinase.
GN M0897 OR MLCB268.15.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1769;
[1]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmier K., Parkhill J., James K.D., Thomson N.R., Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D., Mungall K., Basham D., Brown D., Chillingworth T., Connor R., Davies R.M., Devlin K., Duchoy S., Feltwell T., Fraser A., Hamlin N., Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S., Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M., Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R., Barrell B.G.;
RA "Massive gene decay in the leprosy bacillus.";
RT Nature 409:1007-1011 (2001).
RL -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC
DR EMBL; AL022602; CAA18685.1; -;
DR EMBL; AL583920; CAC31278.1; -;
DR LepToma; M0897; -;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE; PS00113; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase;
KW Complete proteome.
SQ SEQUENCE 400 AA; 43592 MW; 16389D0516EFD21 CRC64;

Query Match 57.5%; Score 45; DB 16; Length 400;
Best Local Similarity 66.7%; Pred. No. 22; Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLARAP 15

DB 86 IEGPTLRQWLARGP 100

RESULT 11

OC 09R9Y9 PRELIMINARY; PRT; 766 AA.
AC 09R9Y9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cti.
GN Cti.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OX NCBI_TaxID=303;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=2440;
RA Junker F., Ramos J.L.;
RA "Involvement of the cis/trans isomerase CtiT1 in solvent resistance.";
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF110739; AAD41255.1; -;
DR InterPro; IPR000345; CytC heme bind.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
SQ SEQUENCE 766 AA; 87108 MW; F6EA2038116239AF CRC64;

Query Match 57.5%; Score 45; DB 2; Length 766;
Best Local Similarity 64.3%; Pred. No. 42; Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EGPTLRWLARAP 15
| : : : : :
Db 177 EYDTRWLARAP 190

RESULT 12

O33466 PRELIMINARY; PRT; 766 AA.
AC O33466;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Cis/trans isomerase.
GN Ctl.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
[1]
SEQUENCE FROM N.A.
RC STRAIN=P8;
RX MEDLINE=98027367; PubMed=9361416;
RA Holtwick R., Meinhardt F., Kewelch H.;
RT "Cis-trans isomerization of unsaturated fatty acids: cloning and
sequencing of the cti gene from Pseudomonas putida P8.";
RL Appl. Environ. Microbiol. 63:4292-4297(1997).
DR EMBL; AJ000978; CAA04438.1; -;
KW Isomerase.
SQ SEQUENCE 766 AA; 86920 MW; 7352654A8321EEFD CRC64;

Query Match 57.5%; Score 46; DB 2; Length 766;
Best Local Similarity 64.3%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EGPTLRWLARAP 15
| : : : : :
Db 177 EYDTRWLARAP 190

RESULT 13

O9JIX1 PRELIMINARY; PRT; 1256 AA.
AC O9JIX1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Nephrlin.
GN NPHS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=20281371; PubMed=10820162;
RA Putaala H., Sainio K., Sariola H., Tryggvason K.;
RT "Primary Structure of Mouse and Rat Nephrlin cDNA and Structure and
Expression of the Mouse Gene.";
RL J. Am. Soc. Nephrol. 11:991-1001(2000).
DR EMBL; AF172256; AAF91087.1; -;
DR MGD; MGI:1859637; Nphs1.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR003598; IG C2.
DR InterPro; IPR003600; IG like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000399; TPP_enzyme.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 8.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00410; IG-like; 3.
DR PROSITE; PS00187; TPP_ENZYMES; UNKNOWN_1.
KW Immunoglobulin domain.

SQ SEQUENCE 1256 AA; 136418 MW; 763706FDD808550F9 CRC64;

Query Match 57.5%; Score 46; DB 11; Length 1256;
Best Local Similarity 46.7%; Pred. No. 69;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 IEPTLRWLARAP 15
| : : : : :
Db 58 VEGSTIKLWGVAP 72

RESULT 14

O8XZR4 PRELIMINARY; PRT; 87 AA.
AC O8XZR4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable molybdopterin MPT converting factor (Subunit 1) protein.
GN MOAD OR RSC1331 OR RS02853.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
[1]
SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
Siguier P., Thébault P., Whalen M., Wincker P., Levy M.,
Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646064; CAD15033.1; -;
DR InterPro; IPR003749; This.
DR Pfam; PF02597; DUF170; 1.
KW Complete proteome.
SQ SEQUENCE 87 AA; 9479 MW; 968734E9BAEF0460 CRC64;

Query Match 56.2%; Score 45; DB 16; Length 87;
Best Local Similarity 80.0%; Pred. No. 7.1;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 LRWLARAP 15
| : : : : :
Db 36 LRWLARAP 45

RESULT 15

O66272 PRELIMINARY; PRT; 245 AA.
AC O66272;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN PUFL.
OS Erythrobacter litoralis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC Erythrobacter.
OX NCBI_TaxID=39960;
[1]
SEQUENCE FROM N.A.
RC STRAIN=IAM14332;
RA Hamada T.;
RT "Use of gyrB gene, pufl and pufM genes and 16S rRNA sequence analysis
to investigate phylogeny of photosynthetic bacteria.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB010981; BAA25791.1; -;
DR HSP; P19057; 1CLT.

DR InterPro; IPR000484; Photo_RC.
 DR Pfam; PF00124; PHOTOC; 1.
 DR PRINTS; PR00256; REACTNCENTRE.
 DR ProDom; PD000551; Photo_RC; 1.
 DR TIGRFAMs; TIGR01157; pufL; 1.
 DR PROSITE; PS00244; REACTION_CENTER; 1.
 FT NON_TER 1
 SQ SEQUENCE 245 AA; 27214 MW; 52B268713E199ABD CRC64;
 Query Match 56.2%; Score 45; DB 2; Length 245;
 Best Local Similarity 80.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 IEQPTLRQWL 10
 |||||
 Db 26 IEQPTLRQWL 35

Search completed: June 24, 2003, 12:52:26
 Job time : 34.4615 secs